PCT / SE 2004 / 0 0 1 5 7 1

29 -10- 2004

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

REC'D 23 NOV 2004
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed R. Mahoney

Dated 25 October 2004

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH
RULE 17.1(a) OR (b)

Patents Form 1/77

Patents Act 1977
(Rule 16)

03NOV03 E848793-1 D00192
P01/7700 0.00-0325494.3

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference — N.90217 GCW

2. Patent application number *(The Patent Office will fill this part in)* — 0325494.3 — 31 OCT 2003

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

SBL VACCIN AB
LUNDAGATAN 2
S-105 21 STOCKHOLM
SWEDEN

Patents ADP number *(if you know it)* 8745127001

If the applicant is a corporate body, give the country/state of its incorporation — SWEDEN

4. Title of the invention — EXPRESSION SYSTEM

5. Name of your agent *(if you have one)* — J. A. KEMP & CO.

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

14 South Square
Gray's Inn
London
WC1R 5JJ

Patents ADP number *(if you know it)* 26001

6. Priority: Complete this section if you are declaring priority from one or more earlier patent applications, filed in the last 12 months.

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. Divisionals, etc: Complete this section only if this application is a divisional application or resulted from an entitlement dispute (see note f)

Number of earlier UK application	Date of filing *(day / month / year)*

8. Is a Patents Form 7/77 (Statement of inventorship and of right to grant of a patent) required in support of this request? — YES

Answer YES if:

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.

Otherwise answer NO (See note d)

Patents Form 1/77

9. Accompanying documents: A patent application must include a description of the invention. Not counting duplicates, please enter the number of pages of each item accompanying this form:

Continuation sheets of this form	-
Description	83
Claim(s)	4
Abstract	1
Drawing(s)	35 + 35

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent (Patents Form 7/77)	
Request for a preliminary examination and search (Patents Form 9/77)	
Request for a substantive examination (Patents Form 10/77)	
Any other documents (please specify)	

11. I/We request the grant of a patent on the basis of this application.

Signature(s) J.A. KEMP & CO. Date 31 OCTOBER 2003

12. Name, daytime telephone number and e-mail address, if any, of person to contact in the United Kingdom

G. C. WOODS
020 7405 3292

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered YES in part 8, a Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) Part 7 should only be completed when a divisional application is being made under section 15(4), or when an application is being made under section 8(3), 12(6) or 37(4) following an entitlement dispute. By completing part 7 you are requesting that this application takes the same filing date as an earlier UK application. If you want the new application to have the same priority date(s) as the earlier UK application, you should also complete part 6 with the priority details.

EXPRESSION SYSTEM

FIELD

The present invention relates to a stable expression system for producing the B subunit of cholera toxin (CTB), a stable expression vector for use in the stable expression system, a stable CTB preparation essentially free of antibiotic residues, vaccines comprising the CTB preparation and uses thereof.

BACKGROUND

The non-toxic B subunit of cholera toxin (CTB) is an effective oral immunising agent, which in a large field trial, has been shown to afford protection against both cholera and enterotoxigenic *E. coli* caused diarrhoea (Sanchez and Holmgren 1989 PNAS 86: 481-485). This has made CTB, as such, an important component, together with killed whole *V. cholerae* cells, of an oral cholera vaccine. Moreover, CTB has attracted much interest recently as an immunogenic carrier for various other peptides or carbohydrate antigens and as an immunomodulator for down regulating the immune response. These findings have emphasised the need to increase the yields of CTB for large scale production to facilitate, in part, vaccine development based on the use of CTB.

The choice of expression system for producing CTB depends on many factors, including the proteolytic stability of the protein, whether or not the protein is secretable and the acceptable costs of the final CTB product. There are four major expression systems which are commonly used to produce vaccine antigens. These are bacterial, yeast, insect and mammalian expression systems. In addition, transgenic plant expression systems have started to emerge with the aim of utilising the plant both for production of the subunit vaccine and for vaccine delivery via the edible plant. By way of example, WO 99/54452 discloses chimeric gene constructs comprising a CTB coding sequence and an autoantigen coding sequence, plant cells and transgenic plants transformed with said chimeric gene constructs, and methods of preparing an edible vaccine from these plant cells and transgenic plants

The expression of recombinant genes in bacterial host cells is most often achieved by the introduction of episomal self-replicating elements (such as plasmids) that encode the structural gene of the protein of interest under the control of an appropriate promoter, into host bacteria. Such plasmids are most commonly maintained by the inclusion of selective marker genes that encode proteins that confer resistance to specific antibiotics (such as ampicillin, chloramphenicol, kanamycin, tetracycline and the like). The plasmids are then maintained in the host by addition of the appropriate antibiotic to the culture medium.

Whilst *E. coli* is the most commonly used bacterium for production of heterologous proteins, the expression of recombinant antigens in bacterial systems other than *E. coli* may sometimes be advantageous. *Salmonella typhimurium*, *V. cholerae* and *Bacillus brevis* are some examples of other bacteria that have been used for expression of antigens for vaccine production purposes.

Known expression systems using *Vibrio cholerae* host cells for the production of heterologous proteins include but are not limited to the CTB expression system disclosed in Sanchez and Holmgren, Proc. Natl. Acad. Sci. USA 1989: 86: 481-5. Details of this expression system are also disclose in U. S. Patent Nos 5268276, 5834246, 6043057 and EP Patent No 0368819. In this expression system, the CTB subunit is obtained by expressing the gene encoding cholera toxin B subunit in a *V. cholerae* host cell in the absence of a *V. cholerae* gene encoding the A subunit of cholera toxin (CTA).

Lebens *et al* (1993 Biotech 11; 1574-8) described a modification of the method of Sanchez and Holmgren (1989 *ibid*) for preparing CTB. In this regard, recombinant CTB was produced by a mutant strain of *V. cholerae* 01, deleted of its CT genes and transfected with a multicopy plasmid encoding CTB. The CTB used was purified from the culture medium by a combination of salt precipitation and chromatographic methods, as described.

The use of bacterial host cell, such as *V. cholerae* host cells for expression of recombinant proteins as demonstrated by Sanchez and Holmgren (1989) (*ibid*) and Lebens *et al* (19930 (*ibid*) has been shown to be advantageous over other prokaryotic expression systems in common use in that specific recombinant products may be produced in large quantities and secreted into the culture medium, thereby facilitating downstream purification procedures. This efficient secretion of CTB from *V. cholerae* host cells is different from the secretory process from *E. coli* cells where the expressed product often assembles in the periplasmic space (Neill *et al* 1983 Science. 221: 289-290). However, recently, a protein secretory pathway for the secretion of heat-labile enterotoxin (LT) by an enterotoxigenic strain of *E. coli* has been identified (Tauschek *et al* (2002) PNAS 99: 7066-7071) which envisage the efficient secretion of a recombinant protein from an *E. coli* host cell.

Whilst the expression system disclosed in Sanchez and Holmgren 1989 (*ibid*) and Lebens *et al* (*ibid*) appear to produce CTB at acceptable levels, these expression systems suffer from the disadvantage that an antibiotic, such as ampicillin, is required in the culture medium to maintain optimum production by selecting for and maintaining plasmids comprising a gene of interest. In the absence of ampicillin, the plasmid containing the gene encoding the CTB subunit protein would not be stably maintained and the yield of the CTB would decrease. In addition, a further downstream processing step is required to effectively remove all the antibiotic residues from the purified product.

The use of antibiotics in the production of recombinant proteins is undesirable for a number of reasons. Apart from the obvious increase in costs arising from the need to add antibiotics as a supplement to the growth medium, the use of antibiotics is considered a problem in the production of any recombinant protein intended for human or veterinary use. This is primarily for three reasons. Firstly, residual antibiotics may cause severe allergic reactions in sensitive individuals. Secondly, there is the possibility of selection for antibiotic resistant bacteria in the natural bacterial flora of those using the product. Finally, DNA encoding the antibiotic

resistance may also be transferred to sensitive bacteria in individuals using the product, thereby also spreading undesired antibiotic resistance in a cohort.

As the large scale production of recombinant proteins, such as CTB, which are free of antibiotic residues, is commercially important in the pharmaceutical industry, there is a need to provide better and safer ways of preparing CTB while maintaining or improving yield. Accordingly, one of the objectives of the present invention is how to provide a safer expression system for preparing CTB while maintaining or improving yield.

SUMMARY OF THE INVENTION

This present invention teaches how to improve CTB yields using a CTB production system comprising a bacterial host cell lacking the functionality of a *thyA* gene which is used in conjunction with a novel expression vector to produce unexpected high yields of CTB relative to the yields obtained with known bacterial host cell production systems.

A plasmid expression vector was constructed in which a gene encoding a thymidylate synthase enzyme (*thyA*) gene was used as a means of selection and maintenance of a plasmid comprising a CTB gene. The plasmid is of reduced size relative to known expression plasmids for producing CTB because substantially all of the non coding *V. cholerae* DNA downstream of the CTB gene is removed.

The unexpected high yield of CTB obtained using this expression system demonstrated both the efficiency of expression of a heterologous gene in a bacterial host cell and the stability of the plasmids maintained by complementation of a *thyA* deletion in the bacterial host cell strain. By way of example, even after repeated passages through liquid culture equivalent to 100 generations all the cells retained the plasmid and the ability to express the recombinant protein.

The expression system as reported here is advantageous because it facilitates the production of CTB for the following uses which include but are not limited to:
a protective immunogen in oral vaccination against cholera and LT-caused *E. coli* diarrhoea;
An immunomodulator or a tolerogenic inducing agent or an immune-deviating agent for down-regulating, modulating, de-sensitising or re-directing the immune response;
An adjuvant for altering, enhancing, directing, re-directing, potentiating or initiating an antigen-specific or non-specific immune response;
A carrier to stimulating an immune response to one or more unrelated antigens; and
A diagnostic agent for producing antibodies (such as monoclonal or polyclonal antibodies) for use in diagnostic or immunodiagnostic tests.
It is a particular advantage from the point of purification and standardisation of CTB as a vaccine component that relatively high yields of CTB can be achieved using stable bacterial host cell strains that lack the functionality of a *thyA* gene.

ASPECTS OF THE INVENTION

In one aspect, the present invention provides a novel and improved stable expression system comprising a combination of (i) a a stable bacterial host cell strain lacking the functionality of a *thyA* gene; and (ii) a stable expression vector comprising a functional *thyA* gene and a CTB gene which is substantially free of the flanking sequences immediately contiguous by the 5' and 3' end of the CTB gene in the naturally occurring genome of the host cell from which the CTB gene is derived..

The stable expression system is advantageous because it provides it;
(i) ensures the stable maintainance of the CTB encoding plasmid (by ensuring, for example, a 100% plasmid retention in the large production fermentor) which is advantageous because it ensures a consistent and reliable production of CTB; and
(ii) improves on CTB quality by eliminating the heterogeneity found in the N-terminus of CTB which ensures consistent production of the same CTB end product.

The invention also provides an isolated stable expression vector for producing CTB which is an improvement over the known expression vectors for producing CTB while still comprising a functional *thyA* gene. The expression plasmid is of reduced size because it eliminates substantially all of the *V. cholerae* DNA downstream of the CTB gene. Without wishing to be bound by theory, it is believed that removing substantially all of non-coding *V. cholerae* DNA downstream of the *ctxB* gene resulting in the reduced size of the expression vector contributes to the improved stability and the improved yield of the CTB product. By way of example of the improved plasmid stability, when *V. cholerae* host cells are used in the expression system, almost all the *V. cholerae* cells retained (i) the plasmid comprising the CTB gene and (ii) the ability to express the recombinant CTB protein even after repeated passage through liquid culture equivalent to 100 generations,

The presence of a functional *thyA* gene in the expression vector is advantageous because:
It complements the *thyA* deficiency in the bacterial host strain;
It enables the strain to grow in the absence of thymine in the growth medium; and
It ensures the genetic stability of the bacterial host strain when grown in a medium devoid of extraneous thymine since loss of the plasmid leads to death of the host strain.

For some embodiments, when the bacterial host cell is a *V. cholerae* host cell, the nucleotide sequence encoding the functional thymidylate synthase (*thyA*) enzyme is an *E. coli* nucleotide sequence or derivable from *E. coli*. The use of plasmid comprising a nucleotide sequence encoding a *thyA* enzyme derivable from *E. coli* is advantageous because the *V. cholerae thyA* gene has only about 30% homology with the corresponding *thyA* sequence from *E. coli* so the risk of a recombination event is reduced.

A further aspect of the invention is a method for producing a cholera toxin B (CTB) subunit protein wherein the method comprises introducing the defined stable expression vector into a bacterial host cell lacking the functionality of a *thyA* gene, and cultivating the host cell under conditions whereby CTB is produced.

The advantages associated with this method include:

(i) improved yield of CTB such that the yield of CTB from the expression system is increased 4-5 fold relative to known CTB expression systems (for example, levels of CTB produced using the known CTB expression systems as described in Sanchez and Holmgren (1989) (*ibid*);

(ii) simplification of the production process for CTB because the downstream step of removing antibiotic residues from CTB can be eliminated. The simplification of the production process results in a cheaper product because there is a reduction in costs in the large scale production of the protein and of the elimination of the need for "down stream processing step" to remove any antibiotic residues from the expressed CTB product.

In another aspect of the invention, there is provided an isolated stably expressed CTB obtainable by the defined expression system wherein the expression system does not express an antibiotic resistance marker and wherein the CTB is essentially free of antibiotic residues.

This stably expressed CTB is advantageous because it is a safer CTB end product by eliminating the antibiotic selection used in strain 213 to maintain the CTB encoding plasmid. The elimination of antibiotic selection means a reduction in the possible incidence of allergic reactions to antibiotics and a reduction in the possible risk of developing antibiotic resistance. Other advantages are discussed and are made apparent by the following commentary

Other aspects of the present invention are presented in the accompanying claims and in the following description and drawings. These aspects are presented under separate section headings. However, it is to be understood that the teachings under each section are not necessarily limited to that particular section heading.

DETAILED DESCRIPTION

Before describing the present invention in detail, it is to be understood that this invention is not limited to particularly exemplified molecules or process parameters as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments of the invention only, and is not intended to be limiting. In addition, the practice of the present invention will employ, unless otherwise indicated, conventional methods of virology, microbiology, molecular biology, recombinant DNA techniques and immunology all of which are within the ordinary skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *DNA Cloning: A Practical Approach*, vol. I & II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *A Practical Guide to Molecular Cloning* (1984); and *Fundamental Virology*, 2nd Edition, vol. I & II (B.N. Fields and D.M. Knipe, eds.).

All publications, patents and patent applications cited herein, whether *supra* or *infra*, are hereby incorporated by reference in their entirety. It must be noted that, as used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the content clearly dictates otherwise.

For the avoidance of doubt, the term "comprising" encompasses "including" as well as "consisting". By way of example, a composition "comprising" X may consist exclusively of X or may include something additional to X such as X and Y.

All scientific and technical terms used in this application have meanings commonly used in the art unless otherwise specified. The standard nomenclautre as used in, for example, E-L Winnacker, From Genes to Clones, VCH Publishers, New York (1987) is adhered to for defining DNA restriction endonucleases, restriction sites and restriction sequences. Oligodeoxynucleotides and amino acids are referred to with the conventional one-letter and three-letter abbreviation codes. The one-letter amino acid symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission are provided at the beginning of the Example Section.
As used in this application, the following words or phrases have the meanings specified.

EXPRESSION SYSTEM
The present invention relates to a stable expression system comprising a host cell and expression vector combination to produce CTB.

The term "expression system" refers to a combination of a host cell and a compatible expression vector which are maintained under suitable conditions, such as, for example, the expression of a protein coded for by foreign DNA carried by the vector and introduced into the host cell. In the case of the expression system as described herein, a *thyA* gene which is essential for bacterial survival is rendered non-functional on the bacterial host cell chromosome. A functional *thyA* gene is provided on a complementing plasmid. The *thyA* gene acts as a selection marker since loss of the plasmid will therefore mean the bacterial host cell is unable to survive. The selection of a *thyA* gene as the non-antibiotic selection marker provides particular advantages as outlined above.

The terms "express" and "expression" includes allowing or causing the information in a gene or DNA sequence to become manifest, for example by producing RNA (such as rRNA or mRNA) or by producing a protein by activating the cellular functions involved in transcription and translation of a corresponding gene or DNA sequence. A DNA sequence is expressed by a cell to form an "expression product" such as an RNA (such as, for example, a mRNA or a rRNA) or a protein. The expression product itself, such as, for example, the resulting RNA or protein, may also said to be "expressed" by the cell.

HOST CELL
As used herein, the term "host cell" refers to any cell of any organism that is selected, modified, transformed, grown or used or manipulated in any way for the production of a substance by the cell. For example, a host cell may be one that is manipulated to express a particular gene, a DNA or RNA sequence, a protein or an enzyme. Host cells can further be used for screening or other assays. The term "host cells" may denote, for example, bacterial cells, that can be, or have been, used as recipients for recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transformed. It is understood that the progeny of a single parental cell

may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation. By way of example, the CTB of the present invention may be expressed in bacterial host cells including but not limited to *E. coli, V. cholerae* or other *Vibrionaceae* bacterial host cells.

In one embodiment, the present invention relates to a CTB production system comprising a *V. cholerae* bacterial host cell lacking the functionality of a *thyA* gene which is used in conjunction with a novel expression vector to produce unexpected high yields of CTB relative to the yields obtained with known bacterial host cell production systems.

In one embodiment, the present invention relates to a CTB production system comprising an *E. coli* bacterial host cell lacking the functionality of a *thyA* gene which is used in conjunction with a novel expression vector to produce unexpected high yields of CTB relative to the yields obtained with known bacterial host cell production systems.

VIBRO GENUS

It is well known in the art that the *Vibrio* genus contains motile, gram-negative bacteria. According to Guidelines for the Epidemiological Investigation of Cholera and other *Vibrio* Species Infections from The Maryland Department of Health and Mental Hygiene Epidemiology and Disease Control Program August 1998, there are 13 known Vibrio subtypes. The most famous of these subtypes is *V. cholerae*. Although all subtypes of *V. cholerae* can cause illness, only disease caused by subtypes 01 and 0139 is considered to be able to be spread epidemically. The thirteen *V. cholerae* subtypes are outlined as follows: *V. cholerae O1, V. cholerae O139, V. cholerae non-O1, non-O139, V. parahaemolyticus, V. mimicus, V. alginolyticus, V. cincinnatiensi, V. fluvialis, V. hollisae, V. vulnificus, V. damsela, V. furnissii*, and *V. metschnikovi.*

VIBRIO CHOLERAE HOST CELLS

It is well known in the art that *V. cholerae* of serogroup O1 and 0139 may induce severe diarrhoeal disease when multiplying in the gut of infected individuals by releasing cholera toxin (CT) which induces active electrolyte and water secretion from the intestinal epithelium. By analogous mechanisms several other bacteria, for instance enterotoxigenic *E. coli* bacteria (ETEC), may also cause diarrhoea by releasing other enterotoxins that may be related or unrelated to CT.

CT is the prototype bacterial enterotoxin. It is a protein built from two types of subunits: a single A subunit of molecular weight 28,000 and five B subunits, each with a molecular weight of 11,600. The B subunits are aggregated in a ring by tight noncovalent bonds; the A subunit is linked to and probably partially inserted in the B pentamer ring through weaker noncovalent interactions. The two types of subunits have different roles in the intoxication process: the B subunits are responsible for cell binding and the A subunit for the direct toxic activity.

The molecular aspects of toxin binding to intestinal and other mammalian cells and of the subsequent events leading to activation of adenylate cyclase through the

intracellular action of the A subunit (and its A1 fragment) have been clarified in considerable detail (see J Holmgren, Nature 292:413-417, 1981). More recently information has also become available on the genetics and biochemistry of cholera toxin synthesis, assembly and secretion by *V. cholerae* bacteria.

CT is encoded by chromosomal structural genes for the A and B subunits, respectively. These genes have been cloned from several strains, and their nucleotide sequences have been determined (see for example, Heidelberg *et al* (2000) Nature 406: 477-483). The genes for the A and B subunits of CT are arranged in a single transcriptional unit with the A cistron (ctxA) preceeding the B cistron (ctxB). Studies on the organization of CT genes in *V. cholerae* strains of classical and El Tor biotypes have suggested that there are two copies of CT genes in classical biotype strains while there is only one copy in most El Tor strains (J J Mekalanos *et al*, Nature 306:551-557, 1983). The synthesis of CT is positively regulated by a gene, toxR that increases ctx expression manifold (V L Miller and J J Mekalanos, Proc Natl Acad Sci USA, 81:3471-3475, 1984). ToxR acts at the transcriptional level, and is present in strains of both classical and El Tor biotypes. ToxR probably increases ctx transcription by encoding a regulatory protein that interacts positively with the ctx promoter region.

E COLI HOST CELLS

Enterotoxigenic *E. coli* (ETEC) bacteria secrete at least one or two types of entertoxins, known as heat-labile (LT) enterotoxin and heat-stable enterotoxin (ST). Although both CTB and LTB proteins may exhibit internal variation in a few amino-acid residues (such as, for example, in human versus animal ETEC isolates and in classical versus El Tor biotype cholera strains), generally, LTB and CTB show a high degree of homology with about 85% conservation of amino acids in the mature protein (Domenighini *et al* (1995) Mol Microbiol 15: 1165-1167) and there is evidence from crystallographic studies that LTB and CTB pentamers are also structurally similar. There is also a high degree of immunological cross-reactivity between the two molecules despite the fact that the majority of antibodies are directed against structural features of assembled pentamers. This is a further indication of the structural similarity between the two molecules.

Studies on heat-labile enterotoxin (LT) in *E coli* (the subunit structure and function of LT is closely similar but not identical to CT) have shown that the A and B subunits are initially synthesized as precursors with a leader peptide preceeding the mature subunit proteins. These precursors are rapidly processed (i.e. the leader peptide is being removed) and translocated across the inner membrane into the periplasm, where unassembled monomeric B subunits pentamerize and associate with A subunit with a half-time of 1-2 min. The pathway of toxin assembly appears to proceed via A subunit association with B monomers or small oligomers. Once the complete toxin has assembled, in *V. cholerae* (in contrast to *E.coli* where the toxin remains in the periplasm) the toxin is being translocated (secreted) across the V. cholerae O1 outer membrane through some sort of interaction of B subunit domains with the outer membrane (T R Hirst & J Holmgren, Proc Natl Acad Sci USA, 84:7418-7422, 1987).

If the B subunits of CT or LT are being expressed in the absence of any A subunit (several such strains have been prepared by chemical mutagenesis or deletions by

recombinant DNA methods in the ctxA or eltA cistrons) the B subunits form pentamers which are then secreted from *V. cholerae* via the same pathway as for the intact toxin except for an apparently slightly slower assembly process in the periplasm (T R Hirst *et al*, Proc Natl Acad Sci USA 81:2645-2649, 1984).

As mentioned above, the efficient secretion of CTB from *V. cholerae* host cells is different from the secretory process from *E. coli* cells where the expressed product often assembles in the periplasmic space (Neill *et al* 1983 Science. 221: 289-290). However, recently, a protein secretory pathway for the secretion of heat-labile enterotoxin (LT) by an enterotoxigenic strain of *E. coli* has been identified (Tauschek *et al* (2002) PNAS 99: 7066-7071). In particular, a novel type II protein secretion pathway has been identified and sequenced from one particular enteropathogenic *E. coli* strain (GenBank Accession No AF426313) (see Tauschek *et al ibid).* Moreover, Horstman and Kuehn (2000) (J Biol Chem 275: 12489-12496) have demonstrated in ETEC strain ETEC 2(ATCC43886) that LT does not remain in the periplasm but is found associated with the cell exterior. Thus, the apparent disadvantages of using an *E. coli* host cell strain to produce the CTB subunit protein maybe overcome by, for example, using a particular ETEC host cell strain (such as ETEC2) or by manipulating the ETEC host cell strain so that it utilises a functional Type II secretory pathway for secreting a CTB subunit protein.

The host cell of the expression system of the present invention is a bacterial host cell. When a bacterial host cell strain is used to produce a protein for use in a vaccine or when the bacterial host cell strain itself is intended itself for use in a vaccine, one of the major concern is that the bacterial host cell vaccine strain is attenuated in the sense that it lacks the ability to synthesise a non-toxic component or form of cholera toxin while still maintaining its immunogenicity.

The bacterial host cell strain as described herein is an attenuated host cell strain. As used herein, the term "attenuated vaccine strain" refers to a vaccine strain which is rendered avirulent so it is non-toxic or so that it does not produce either a toxic components or there is no risk that it will revert to toxigenicity which could cause disease.

An attenuated bacterial host cell strain may be produced genetically by functionally inactivating a toxic gene in the host cell strain. This functional inactivation may be achieved by methods which include but which are not limited to abolishing or preventing synthesis of any toxic polypeptide by deleting either the entire gene encoding the toxic polypeptide or the functional part of the gene or by making a mutation, such as a deletion or insertion within the coding sequence of a gene that results in synthesis of non-functional polypeptide (such as, for example, a polypeptide that contains only the N-terminal sequence of the wild-type protein). By way of example, mutations may be introduced into into a bacterial strain host cell strain to prevent expression of toxic enterotoxins or other virulence genes or by deleteting or inactivating the relevant enterotoxin or virulence gene. When a mutation includes insertions and/or deletions, then typically, the insertions and/or deletions are preferably large, typically at least 10 nucleotides in length up to the length of the entire gene or coding sequence, for example from 10 to 600 nucleotides. Preferably,

the whole coding sequence or whole gene is deleted. The introduced mutations are generally non-reverting mutations. That is, these mutations that show essentially no reversion back to the wild-type.

In one embodiment, the host cell comprises one or more amino acid additions, substitutions or deletions in the amino acid sequence of the A subunit of CT or LT which is or are effective to abolish the toxicity of the toxin and which results in a detoxified host cell.

As is well known, the toxicity of the cholera toxin (CT) resides in the A subunit and a number of mutants of CT and its homologue, LT, comprising point mutations in the A subunit are known in the art. By way of example WO 92/19265 disclose mutations in the CTA subunit at Arg-7, Asp-9, Arg-11, His-44, His-70 and Glu-112. WO 93/13202 relates to immunogenic detoxified CT and LT proteins having substitutions at one or more of amino acids Val-53, Ser-63, Val-97, Tyr-104 or Pro-106. WO 95/17211 discloses an LT mutant with a Lys-7 mutation (LT-K7). Methods for the design and production of mutants of CT and its homologues, such as, for example, site-directed mutagenesis of DNA encoding the wild type toxins, are known in the art. By way of example, suitable methods are described in WO 93/13202, WO 92/19265, WO 95/17211 and WO 98/42375 the disclosures of which is incorporated herein by reference.

The bacterial host cell as described herein is a detoxified host cell which has reduced virulence such that, it does not cause a bacterial associated diease, such as diarrhoea. As used herein, the term "detoxified host cell" means that the host cell exhibits a substantially lower toxicity relative to its naturally occurring counterpart. The substantially lower toxicity should be sufficiently low for the host cell to be used as part of an expression system or for the host cell itself to be used in an immunologically effective amount as a vaccine without causing significant side effects. The toxicity may be measured in mouse CHO cells or preferably by evaluation of the morphological changes induced in Y1 cells.

In one embodiment, the *V. cholerae* vaccine strains as described herein may be an avirulent or attenuated or a detoxified *V. cholerae* host cell strain comprising a non-functional CTA gene in the *V. cholerae* chromosome. The attenuated *V. cholerae* strain may arise from a site-directed mutagenesis which results from a deletion in both its chromosomal *ctxA* loci thereby abrogating the production of intact holo-toxin or from a disruption or inactivation of ADP-ribosyltransferase activity resulting in a non-toxic holotoxin. This avirulent or attenuated or detoxified *V. cholerae* host cell strain may be further modified so that it lacks the functionality of a *thyA* gene.

A *V. cholerae* host cell strain lacking the functionality of a *thyA* gene can be prepared by methods of the invention or by methods known to those skilled in the art (Sambrook, J. E. F. Fritsch, and T. Maniatis, Molecular cloning : a laboratory manual. 2nd ed. 1989: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y). Appropriate known methods for preparing the *V. cholerae* strain lacking the functionality of the *thyA* gene include the methods outlined in WO 99/61634.

In the context of the present invention, a *thyA* gene lacks functionality if, for example, the gene has been removed – such as by deletion- or if the gene has been genetically disabled by, for example, inactivation or site directed mutagenesis of the *thyA* gene so that there is no expression of the *thyA* enzyme. The lack of functionality of a *thyA* gene may be determined, for example, by transforming a *thyA* negative vector with a *thyA* positive gene and selecting for absence of growth in the absence of thymine.

THYA SELECTABLE MARKER SYSTEM

The complementation of a chromosomal lesion on a bacterial host cell strain has been used as means of plasmid maintenance. Thus, the non-functional *thyA* gene on the *V. cholerae* chromosome is complemented by the presence of a functional *thyA* gene provided on a complementary plasmid which acts as a selectable marker and which eliminates the need to an antibiotic resistance selection marker. The *thyA* gene also acts as a selectable marker in the sense that loss of the plasmid means that the *V. cholerae* bacterium is unable to survive.

The thymidylate synthetase (*thyA*) enzyme encoded by the *thyA* gene of *V. cholerae*, *E.coli* and other bacteria catalyses the methylation of deoxyuridylate (dUMP) to deoxythymidylate (dTMP) and is an essential enzyme in the biosynthesis of deoxyribothymidine triphosphate (dTTP) for incorporation into DNA. In the absence of this enzyme the bacteria become dependent upon an external source of thymine which is incorporated into dTTP by a salvage pathway encoded by the *deo* genes (Milton *et al* 1992 J Bacteriol 174: 7235-7244).

It is known that the *thyA* gene is a conserved gene and can be found in bacteriophages, prokaryotes and eukaryotes. Because the *thyA* enzyme is conserved, the *thyA* gene, from, for example, *E. coli*, is able to complement the mutant *thyA* genes located in the chromosome of related bacterial species as discussed below. The functional *thyA* gene in the plasmid may be from sources other than *E. coli*. In one embodiment, wherein the host cell is a *V. cholerae* host cell, the *thyA* gene has a low homology with the *V. cholerae thyA* gene.

Previous work in the field has demonstrated that recombinant plasmids can be maintained in *V. cholerae* in the absence of antibiotic selection by complementation of a *thyA* mutation with the *thyA* gene from *E. coli*. The principle was demonstrated initially using plasmids carrying the *E. coli thyA* gene and spontaneous *thyA* mutants of *V. cholerae* isolated on the basis of resistance to trimethoprim (Morona *et al* 1991 Gene 107: 139-144).

Further work by Carlin and co-workers resulted in the cloning and characterisation of the *thyA* locus from *V. cholerae* and the generation of stable defined recipient *V. cholerae* strains. The sequence of the *V. cholerae thyA* gene as determined by Carlin and co-workers is published in EMBL (Genebank Accession No AJ006514). WO 99/61634 teaches that defined *thyA* mutants of *V. cholerae* may be used as suitable production strains for recombinant proteins encoded on plasmids maintained by *thyA* complementation.

The use of a *thyA* gene on the complementing plasmid, which has low homology with the *V. cholerae thyA* gene is advantageous because the risk of "cross-over" with the *V. cholerae* chromosome is reduced.

Preferably the *thyA* gene on the complementing plasmid is an *E. coli thyA* gene which has a low homology with *V. cholerae thyA* gene. By way of explanation, the published sequence for the *E. coli thyA* gene can be found at Genebank Accession No J01709. A comparison of the sequence of the *V. cholerae thyA* gene (protein of 283 amino acids) as determined by Carlin *et al* (see Genebank Accession No AJ006514) and the *E. coli thyA* gene (see Genebank Accession No J01709) showed only 32% amino acid identity and reflects only about 54% homology in a 454bp overlap at the DNA level (see Figure 7 of WO 99/61634). In this regard, homology searches of the EMBL DNA and Swiss-Prot protein data libraries were done by the FASTA software in the GCG program package (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI.)

In another embodiment of the present invention, an *E. coli* host cell lacking the functionality of a ThyA gene is complemented by a plasmid with a functional ThyA gene, preferably a *V. cholerae* ThyA gene.

BACTERIAL EXPRESSION SYSTEM

The expression system of the present invention comprises a host cell and an expression vector to produce CTB. When a bacterial host cell is used, the expression system is typically referred to as a bacterial expression system or a bacterial host cell expression system. Bacterial expression systems are known in the art. By way of example, bacterial expression systems may comprise one or more expression control elements, such as but not limited to: a promoter, an enhancer, a transcription initiation site, a ribosome binding site (RBS), a leader sequence (signal sequence), a translational stop codon, and a transcription terminator sequence, all operably linked to the gene or genes which need to be expressed (such as the CTB gene). The term "control sequences" is intended to include, at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

PROMOTER

As used herein, a "promoter sequence" refers to a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3'direction) coding sequence into mRNA. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently found, for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps initiate transcription of the lac operon in *E.coli* [Raibaud *et al.* (1984) Annu. Rev. Genet. 18:173]. Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (lac) [Chang *et al.* (1977) Nature 198:1056], and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (trp) [Goeddel *et al.* (1980) Nuc. Acids Res. 8:4057; Yelverton *et al.* (1981) Nucl. Acids Res. 9:731; U.S. Pat. No. 4,738,921; EPO Publ. Nos. 036 776 and 121 775]. The g-lactamase (bla) promoter system [Weissmann (1981) "The cloning of interferon and other mistakes." In Interferon 3 (ed. I. Gresser)], bacteriophage lambda PL [Shimatake *et al.* (1981) Nature 292:128] and T5 [U.S. Pat. No. 4,689,406] promoter systems also provide useful promoter sequences.
In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter [U.S. Pat. No. 4,551,433]. For example, the tac promoter (tacP) is a hybrid trp-lac promoter comprised of both tri promoter and lac operon sequences that is regulated by the lac repressor [Amann *et al.* (1983) Gene 25:167; de Boer *et al.* (1983) Proc. Natl. Acad. Sci. 80:21]. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system [Studier *et al.* (1986) J. Mol. Biol. 189:113; Tabor *et al.* (1985) Proc Natl. Acad. Sci. 82:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E.coli* operator region (EPO Publ. No. 267 851).

The expression of CTB as described herein may be driven by a variety of promoters. Preferably the promoter is a heterologous promoter. As used herein, the term "heterologous" refers to two biological components that are not found together in nature. The components may be regulatory regions, such as promoters. As used herein, the term "heterologous promoter" refers to a promoter which is unrelated to the gene with which is it operably linked. Preferably the promoter is a heterologous prokaryotic promoter. In particular, preferably the promoter is a promoter suitable for the host cell in which it will be used. More preferably, expression of the CTB gene in

the expression system as described herein is driven by the tacP Promoter or T7 RNA polymerase dependent promoter. In one embodiment, CTB may be expressed in an inducible (such that a stimulus is required to initiate expression) or a constitutive manner (such that it is continually produced) under the control of a heterologous promoter, such as the tacP promoter. In the case of inducible expression, the production of rCTB can be initiated when required by, for example, addition of an inducer substance to the culture medium, for example dexamethasone or Isopropylthiogalactoside (IPTG) which is an artificial inducer of the Lac operon.

RIBOSOMAL BINDING SITE

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E.coli,* the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon [Shine *et al.* (1975) Nature 254:34]. The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' and of *E.coli* 16S rRNA [Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In Biological Regulation and Development: Gene Expression (ed. R. F. Goldberger)]. To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site [Sambrook *et al.* (1989) "Expression of cloned genes in Escherichia coli." In Molecular Cloning: A Laboratory Manual].

TRANSCRIPTION TERMINATION SEQUENCES

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA.

Any suitable transcriptional termination sequence may be used, preferably a strong transcriptional termination sequence which allows minimal or no transcription. In a preferred embodiment of the invention, TrpA terminators are located downstream of the CTB gene effectively terminating mRNA transcription. In one embodiment described in the Examples, the nucleotide sequences of the transcription terminator sequences are shown in Figure 16 (from about nucleotide 2732 to about nucleotide 2759).

LEADER SEQUENCES/SIGNAL SEQUENCES

In one embodiment of the present invention, a DNA molecule encoding CTB may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. In one embodiment, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* on *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

In another embodiment of the present invention, a DNA molecule may be expressed extracellularly but using a signal/leader sequence to ensure secretion of the desired

CTB protein to the culture medium. Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal/leader peptide sequence fragment that provides for secretion of the foreign protein in bacteria [see for example U.S. Pat. No. 4,336,336]. The signal/leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (such as, for example, for gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (such as, for example, for gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E.coli* outer membrane protein gene (ompA) [Masui *et al.* (1983), in: Experimental Manipulation of Gene Expression; Ghrayeb *et al.* (1984) EMBO J. 3:2437] and the *E.coli* alkaline phosphatase signal sequence (phoA) [Oka *et al.* (1985) Proc. Natl. Acad. Sci. 82:7212]. As an additional example, the signal sequence of the alpha-amylase gene from various *Bacillus* strains can be used to secrete heterologous proteins from B. *subtilis* [Palva *et al.* (1982) Proc. Natl. Acad. Sci. USA 79:5582; EPO Publ. No. 244 042].

The CTB produced by the host cell may be secreted or may be contained intracellularly depending on the type of host cell used. As will be understood by those of skill in the art, expression vectors as described herein can be designed with signal sequences/leader sequence which direct secretion of the CTB product through a particular prokaryotic or eukaryotic cell membrane. The CTB protein may be expressed and secreted as an N-terminal extended fusion protein. The N-terminal extension may then be removed from the recovered CTB protein *in vitro* by either chemical or enzymatic cleavage as well known in the art. It is preferred to conduct the cleavage by use of an enzyme.

As used herein, the term "leader sequence" or "signal sequence" relates to any nucleotide encoding sequence or encoded peptide sequence on a protein molecule which facilitates the translocation or exportation of a protein, such as the translocation or exportation of an expressed CTB protein across the cellular membrane and cell wall, if present, or at least through the cellular membrane into the periplasmic space of a cell having a cell wall. As used herein, the term "leader sequence" or "signal sequence" refers to a DNA sequence that encodes a polypeptide (a"secretory peptide") that, as a component of a larger polypeptide or propeptide sequence, which directs the larger polypeptide through a secretory pathway of a cell (such as from the endoplasmatic reticulum to the Golgi apparatus and further to a secretory vesicle) in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway. The secretory signal sequence may encode any signal peptide which ensures efficient direction of the

expressed polypeptide into the secretory pathway of the cell. The signal peptide may be naturally occurring signal peptide, or a functional part thereof, or it may be a synthetic peptide.

In one embodiment, the leader sequence is from an enterotoxin, such as an *E. coli* heat-labile enterotoxin (LT) leader sequence. Examples of LT leader sequences are provided in the sequences listed in Table 1 under their listed G1 Accession numbers. In one preferred embodiment, the leader sequence is *E. coli* heat-labile enterotoxin (LTB) leader sequence). The LTB signal sequence for producing CTB of the present invention is presented in Table 2 as MNKVKFYVLFTA LLSS LCAH. Other examples of leader sequences include but are not limited to leader sequences presented in Table 2 and as part of the sequences presented in Figure 16.

In the described examples, the CTB gene is fused to the LTB signal peptide from the heat-labile enterotoxin of *E. coli* in such a way that the naturally occurring SacI site can be used.

DNA CONSTRUCT

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, may also be called a "DNA construct" or a "nucleic acid construct". Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a procaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign expressed protein on the host cell. Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

REPLICON

As used herein, a "replicon" is any genetic element, e.g., a plasmid, a chromosome, a virus, a cosmid, and the like that behaves as an autonomous unit of polynucleotide replication within a cell. The replicaon is capable of replication under its own control and may include selectable markers.

VECTOR

As used herein, a "vector" is a replicon in which another polynucleotide segment is attached, so as to bring about the replication and/or expression of the attached

segment. The term "vector" includes expression vectors and/or transformation vectors. The term "expression vector" means a construct capable of *in vivo or in vitro/ex vivo* expression. The term "transformation vector" means a construct capable of being transferred from one species to another. Examples of vectors include but are not limited to plasmids, chromosomes, artificial chromosomes or viruses.

PLASMIDS

A common type of vector is a "plasmid", which generally is a self-contained molecule of double-stranded DNA, usually of bacterial origin, that can readily accept additional (such as heterologous) DNA and which can readily introduced into a suitable host cell. A large number of vectors, including plasmid and fungal vectors, have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts. The plasmid employed in the invention may be a plasmid known in the art such as but not limited to plasmids such as pBR322, pACYC177 or pUC plasmid derivatives or the pBLUESCRIPT vector (Stratagene, La Jolla, CA).

Plasmids such as pJS162 (as described in Sanchez and Holmgren (1989) (*ibid*) and (pML358) (as described in Lebens *et al* 1993 *ibid*) have been used to produce CTB in a *V. cholerae* host cell expression systems. The expression vector of the present invention is different from the expression plasmids of Sanchez-Holmgren (pJS162) and Lebens (pML358) in that:

(i) the plasmid is of a smaller size because substantially all of the non-coding *V. cholerae* DNA downstream of the CTB gene has been removed; and

(ii) the plasmid has a functional *thyA* gene.

In this regard, Table 3 provides a comparative analysis of the expression vector as described herein with the relevant expression vectors known in the art. In one embodiment, the stable expression vector as described herein is preferably less than 5kb in size. In a more preferred embodiment, the stable expression vector is from about 2.5kb to 4kb in size. In an even more preferred embodiment, the stable expression vector is about 3kb in size.

The term "about" or "approximately" means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined within the limitations of the measurement system. For example, "about" can mean within 1 or more than 1 standard deviations, as per the practice in the art. Alternatively, "about" can mean a range of up to 20%, preferably up to 10%, more preferably up to 5%, and more preferably still up to 1 % of a given value. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude, preferably within 5-fold, and more preferably within 2-fold, of a value.

The smaller plasmid size is advantageous because it allows easier *in vitro* manipulation and construction of derivatives because smaller DNA molecules ligate together and transform into prokaryotic hosts, such as *V. cholerae*, more efficiently, improving the chances of obtaining derivatives of the correct construction. The smaller size also allows greater efficiency when introducing the constructs into

recipient bacteria by, for example, transformation and also to increase the stability of the plasmid.

EXPRESSION VECTOR

As used herein, the term "expression vector" means the vehicle by which a nucleotide sequence (such as, a heterologous nucleotide sequence) can be introduced into a host cell so as to transform the host and promote expression (such as, for example, transcription and translation) of the introduced sequence.

ISOLATED EXPRESSION VECTOR

As used herein, the term "expression vector" includes an isolated expression vector as well as an expression vector which is part of a host cell/expression vector combination. The terms "isolated" and "purified" refer to molecules, either nucleic or amino acid sequences or nucleic acid constructs that are removed from their natural environment and/or isolated or separated from at least one other component with which they are naturally associated. By way of example, an expression vector may be regarded as "isolated" if it has been prepared under conditions that reduce or eliminate the presence of unrelated materials, such as, for example, contaminants, including native materials from which the material is obtained. By way of further example, a purified protein is regarded as isolated if it is substantially free of other proteins or nucleic acids with which it is associated in a cell. Likewise, a purified nucleic acid molecule is isolated if it is substantially free of proteins or other unrelated nucleic acid molecules with which it can be found within a cell. A protein may be mixed with carriers or diluents which will not interfere with the intended purpose of the substance and still be regarded as substantially isolated.

HETEROLOGOUS NUCLEOTIDE SEQUENCE

Generally, a heterologous nucleotide sequence is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. As used herein, the term "heterologous nucleotide sequence" refers to a nucleotide sequence which is not naturally located in a cell or in a chromosomal site of a cell or which is not naturally expressed by a cell. As used herein, x is "heterologous" with respect to y if x is not naturally associated with y in the identical manner; i.e., x is not associated with y in nature or x is not associated with y in the same manner as is found in nature. The term "heterologous nucleotide sequence" is used interchangeably with the terms "foreign" nucleotide sequence or "guest" nucleotide sequence or "extracellular" nucleotide sequence or "extrinsic" or "exogenous" nucleotide sequence throughout the text. The heterologous nucleotide sequence may also be a coding sequence.

As used herein, the terms "gene", "coding sequence" or a nucleotide sequence "encoding" an expression product, such as a RNA, polypeptide, protein or enzyme, refers to a nucleotide sequence which when expressed, results in the production of that RNA, polypeptide, protein or enzyme. That is, the nucleotide sequence "encodes" that RNA or it encodes the amino acid sequence for that polypeptide, protein or enzyme. The term "nucleotide sequence" is synonymous with the term "polynucleotide". A gene sequence or nucleotide sequence is "under the control of" or is "operably linked with" transcriptional and translational control sequences in a

cell when RNA polymerase transcribes the coding sequence into RNA, which is then trans-RNA spliced (if it contains introns) and, if the sequence encodes a protein, is translated into that protein. The nucleotide sequence may be DNA or RNA of genomic or synthetic or of recombinant origin. The nucleotide sequence may be double-stranded or single-stranded whether representing the sense or antisense strand or combinations thereof.

CODING SEQUENCE

As used herein, a "coding sequence" is a polynucleotide sequence which is translated into a polypeptide, usually via mRNA, when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, cDNA, and recombinant polynucleotide sequences. An "open reading frame" (ORF) is a region of a polynucleotide sequence which encodes a polypeptide; this region may represent a portion of a coding sequence or a total coding sequence.

OPERABLY LINKED

A control sequence may be operably linked to a coding sequence. As used herein, the term "operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

CHOLERA TOXIN (CT) AND B SUBUNIT THEREOF (CTB)

As used herein, the term "CT" refers to the cholera toxin and "CTB" refers to the B subunit of the cholera toxin. In other texts, these may sometimes be identified as "CT" or "Ct" and "CtxB" or "CtB" respectively. The CTB produced by the expression system of the present invention may also be referred to as recombinant CTB (rCTB). The term "CTB" also includes recombinant CTB DNA sequences which are part of a hybrid CTB gene or derivative thereof encoding additional sequences. A CTB derivative could be a fusion protein such as a CTB gene fusion protein or a CTB coupled with other elements.

HEAT-LABILE ENTEROTOXIN (LT) AND B SUBUNIT THEREOF (LTB)

As used herein, the term "LT" herein refers to the *E. coli* heat labile enterotoxin, and "LTB" is the B subunit of LT. In other texts, these may sometimes be identified as "Etx" or "Et" and "EtB" or "EtxB" respectively. The heat-labile toxin (LT) of enterotoxigenic *E. coli* (ETEC) is structurally, functionally and immunologically similar to CTB. The two toxins cross-react immunologically.

CTB GENE

The CTB gene or nucleotide sequence encoding CTB is substantially free from the flanking sequences immediately contiguous by the 5' and 3' end of the CTB encoding sequence in the naturally occurring genome of the micro-organisms from which the CTB encoding DNA is derived. In other words, the CTB gene is substantially free of the 5' and 3' flanking sequences homologous to its host cell genome. For some

applications, the CTB gene or the nucleotide sequence encoding the CTB protein may be the same as the naturally occurring or native form or wild type form of CTB.

"NATIVE CTB"
As used herein the term "native CTB" refers to a CTB molecule with properties, such as activity (such as, for example, GM-1 binding activity) and/or immunogenic and/or immunomodulatory properties which are substantially the same as the naturally occurring form or wild type form of the CTB molecule which is capable of binding to GM1 and/or which have the immunogenic or immunomodulatory capability of the CTB molecule. The terms "native", "naturally occurring", "wild-type" form of CTB are used inter-changeably throughout the text.

In one embodiment (as described in the Examples below), the substantially pure CTB gene is presented as as the nucleotide sequence from about nucleotides 2402 to about nucleotides 2710 in Figure 16.

For some applications, the CTB gene or the nucleotide sequence encoding the CTB protein may be a variants, homologues, derivatives or fragments thereof of the naturally occurring or native form of the CTB.

As used herein, the term "variants, homologues, derivatives and fragments thereof" of a native CTB molecule include CTB molecules which may be structurally different from the native CTB molecule (such as, for example, in terms of nucleotide sequence) but which behave functionally like the native CTB molecule particularly in terms of its binding properties, such as binding to GM1 ganglioside and/or its immunological properties such as reacting with antiserum to CTB as detected by an ELISA or GM1-ELISA test. These variants, homologues, derivatives and fragments thereof of a native CTB molecule include but are not limited to the B subunit of heat-labile enterotoxin from *E. coli* B (LTB) and to any or all mutated, extended, truncated or otherwise modified forms of B subunits or any other protein that would react with GM1 or with said types of antisera as well as any nucleic acid preparation that would encode for a protein that would meet these criteria but which do not have any ADP-ribosylating activity.

In another embodiment, the CTB gene is presented as a variant, homologue, derivative or fragment of the sequence presented from about nucleotide 2402 to about nucleotide 2710 in Figure 16.

"MATURE CTB"
As used herein, the term "mature CTB" refers to the expressed CTB subunit protein which is devoid of a signal sequence. By way of example, the amino acid sequence of mature CTB and mature LTB are depicted in Figure 28 and start with the amino acid +1. The corresponding genes encoding mature CTB and LTB are depicted in the same Figure.

As used herein, the term "amino acid sequence" refers to peptide, polypeptide sequences, protein sequences or portions thereof.

As used herein, the term "protein" is synonymous with the term "amino acid sequence" and/or the term "polypeptide". In some instances, the term "amino acid sequence" is synonymous with the term "peptide". In some instances, the term "amino acid sequence" is synonymous with the term "protein".

As used herein, the term "polypeptide" refers to a polymer of amino acids and does not refer to a specific length of the product. Thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not refer to or exclude post expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like. Included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

A polypeptide or amino acid sequence "derived from" a designated nucleic acid sequence refers to a polypeptide having an amino acid sequence identical to that of a polypeptide encoded in the sequence, or a portion thereof wherein the portion consists of at least 3-5 amino acids, and more preferably at least 8-10 amino acids, and even more preferably at least 11-15 amino acids, or which is immunologically identifiable with a polypeptide encoded in the sequence. This terminology also includes a polypeptide expressed from a designated nucleic acid sequence.

N-TERMINAL MUTATION

Threonine (T) is the first amino acid normally found in mature CTB from the native or naturally occurring form of the CTB molecule. Thus, generally, the N-terminal sequence for mature CTB molecule is Thr-Pro-Gln-Asn-Ile-Thr (TPQNIT). Examples of mature CTB molecules with a TPQNIT N-terminal sequence include but are not limited to the CTB amino acid sequence from *V. cholerae strain* 0395, classical Ogawa, which is shown in US patents Nos 5268276, 58234246 and 6043057, EP Patent No 0368819 and Figure 2 of Sanchez and Holmgren (1989) (*ibid*). The described embodiments (see Figure 15) provide an example of a CTB sequence with a TPQNIT N-terminal sequence which is produced using a *V. cholerae* host cell expression system.

In one embodiment, variants of the CTB sequence may be used which advantageously have the APQNIT (Ala-Pro-Gln-Asn-Ile-Thr) N-terminal sequence. By way of example, the CTB sequence presented in Figure 16 is the same as the CTB native sequence from *V. cholerae strain* 0395, classical Ogawa, apart from the single mutation at the amino terminal end of the protein sequence where an alanine (Ala) residue is introduced at the first position of the CTB amino acid sequence instead of a Threonine (Thr = T) Introduction of this particular amino acid (Ala) is advantageous because it creates a defined signal sequence cleavage site, as opposed to the threonine (Thr) residue at the amino terminus of the wild type or native form of CTB. This cleavage site can be important in post-translational modifications. This N-terminal mutation is advantageous because it improves on CTB quality by eliminating the heterogeneity found in the N-terminus of CTB produced using known CTB expression systems (such as the CTB expression system described in Sanchez and

Holmgren 1989 (*ibid*) and so ensures consistent production of the same CTB end product. In this respect, the junction of the eltBlctxB gene has been modified so that only a single N-terminal is obtained in the resulting CTB protein, in comparison up to about two different N-termini which are obtained with the native CTB molecule (see US patents Nos 5268276, 58234246 and 6043057, EP Patent No 0368819 and Sanchez and Holmgren (1989) (*ibid*).

VARIANT

As used herein, the term "variant" may also be used to indicate a modified or altered gene, DNA sequence, RNA, enzyme, cell and the like which differs from the native type sequence. A variant may be found within the same bacterial strain or may be found within a different strains. Preferably the variant has at least 90% sequence identity with the native or naturally occurring form of the CTB sequence. Preferably the variant has 20 mutations or less over the whole native sequence. More preferably the variant has 10 mutations or less, most preferably 5 mutations or less over the whole native CTB sequence.

MUTANT

As used herein, the terms "mutant" and "mutation" refers to any detectable change in genetic material, such as, for example, any DNA, or any process, mechanism or result of such a change. This includes gene mutations, in which the structure (such as the DNA sequence) of a gene is altered, any gene or DNA arising from any mutation process, and any expression product (such as, for example, RNA, protein or enzyme) expressed by a modified gene or DNA sequence. A mutant may arise naturally, or may be created artificially (such as, for example, by site-directed mutagenesis). Preferably the mutant has at least 90% sequence identity with the native or naturally occurring or wild type CTB sequence. Preferably the mutant has 20 mutations or less over the whole wild-type CTB sequence. More preferably the mutant has 10 mutations or less, most preferably 5 mutations or less over the whole wildtype CTB sequence.

By way of example, a CTB variant may include any subunit protein including at least one mutation, addition, or deletion of residues between positions 1-103 of CTB is disclosed. Examples of such mutations include any point mutation, deletion or insertion into these toxins, subunits or other proteins as well as any peptide extensions to these proteins whether placed in the amino-end, the carboxy-end or elsewhere in the protein and irrespective of whether these peptides have immunological properties by being B cell epitopes, T cell epitopes or otherwise which are capable of stimulating or deviating the immune response. For example, a number of such mutants have been described in the literature (Backstrom *et al*; Gene 1995; 165: 163-171; Backstrom *et al*., Gene 1996; 169: 211-217; Schodel *et al*., Gene 1991; 99: 255-259; Dertzbaugh *et al*. Infect. Immun. 1990; 58: 70-79).

HOMOLOGY

As used herein, the term "homology" refers to the degree of similarity between x and y. The correspondence between the sequence from one form to another can be determined by techniques known in the art. For example, they can be determined by a direct comparison of the sequence information of the polynucleotide. Alternatively,

homology can be determined by hybridization of the polynucleotides under conditions which form stable duplexes between homologous regions (for example, those which would be used prior to S1 digestion), followed by digestion with single-stranded specific nuclease(s), followed by size determination of the digested fragments.

HOMOLOGUE

Any CTB sequences, such as but not limited to that presented in Figure 16 or those described under their GI Accession Numbers in Table 1 may be useful in the present invention. In one embodiment, a CTB protein expressed by a *V. cholerae* host cell of the invention may be encoded by:

(i) a DNA molecule comprising the nucleotide sequence of the CTB gene presented in Figure 16 or specified in Table 1 by GenBank accession number;

(ii) a DNA molecule which hybridises to the complement of the nucleotide sequence in (a); or

(iii) a DNA molecule which encodes the same amino acid sequence as the DNA molecule of (a) or (b) but which is a degenerate form of the DNA molecule of (a) or (b).

As defined herein, the term "homologue" refers to an entity having a certain homology with the native or wild type amino acid sequence and the native or wild type nucleotide sequence. Here, the term "homology" can be equated with "identity". A homologue of the polynucleotide sequence in (i) may be used in the invention. Typically, a homologue has at least 40% sequence identity to the corresponding specified sequence, preferably at least 60%, 70%, 75%, 80% or 85% and more preferably at least 90%, 95% or 99% sequence identity. Such sequence identity may exist over a region of at least 15, preferably at least 30, for instance at least 40, 60 or 100 or more contiguous nucleotides. Typically, the homologues will comprise the same-active-sites and the like as the subject amino acid sequence. Although homology can also be considered in terms of similarity (that is, amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

Methods of measuring polynucleotide homology are well known in the art. Homology comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate percent homology between two or more sequences. Percent homology may be calculated over contiguous sequences. That is, one sequence is aligned with the other sequence and each amino acid in one sequence is directly compared with the corresponding amino acid in the other sequence, one residue at a time. This is called an "ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues.

Although this method is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in percent homology when a

global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalising unduly the overall homology score. This is achieved by inserting "gaps" in the sequence alignment to try to maximise local homology.

However, these more complex methods assign "gap penalties" to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible reflecting higher relatedness between the two compared sequences-will achieve a higher score than one with many gaps. "Affine gap costs" are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons. For example when using the GCG Wisconsin Bestfit package the default gap penalty for amino acid sequences is-12 for a gap and -4 for each extension.

Calculation of maximum percent homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U. S. A.; Devereux *et al* 1984, Nucleic Acids Research 12: 387-395). Examples of other software than can perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel *et al* 1999 *ibid* Chapter 18), FASTA (Atschul *et al* 1990, J. Mol. Biol., 403-410) and the GENEWORKS suite of comparison tools. Both BLAST and FASTA are available for offline and online searching (see Ausubel *et al* 1999 *ibid*, pages 7-58 to 7-60) and Altschul (1993) J Mol Evol 36: 290-300 or Altschul *et al* (1990) J Mol Biol 215: 403-10. However, for some applications, it is preferred to use the GCG Bestfit program. A new tool, called BLAST 2 Sequences is also available for comparing protein and nucleotide sequence (see FEMS Microbiol Lett 1999 174 (2): 247-50; FEMS Microbiol Lett 1999 177 (1) : 187-8 and tatiana@ncbi. nlm. nih. gov).

Although the final percent homology can be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix-the default matrix for the BLAST suite of programs. GCG Wisconsin programs generally use either the public default values or a custom symbol comparison table if supplied (see user manual for further details). For some applications, it is preferred to use the public default values for the GCG package, or in the case of other software, the default matrix, such as BLOSUM62. Once the software has produced an optimal alignment, it is possible to calculate percent homology, preferably percent sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

The homologue may differ from the corresponding specified sequence by at least 1, 2,5, 10 or more substitutions, deletions or insertions over a region of at least 30, for instance at least 40, 60 or 100 or more contiguous nucleotides, of the homologue. Thus, the homologue may differ from the corresponding specified sequence by at least 1,2, 5,10, 30 or more substitutions, deletions or insertions. A homologue CTB gene may be tested by expressing the gene in a suitable host and testing for cross reactivity with antibody specific to the particular CTB antigen.

The expression plasmid used in the present invention may comprise nucleotide sequences that can hybridise to the nucleotide sequences presented herein (including complementary sequences of those presented herein). A homologue typically hybridises with the corresponding specified sequence at a level significantly above background. The signal level generated by the interaction between the homologue and the specified sequence is typically at least 10 fold, preferably at least 100 fold, as intense as background hybridisation. The intensity of interaction may be measured, for example, by radiolabelling the probe, such as, for example, with ^{32}P.

Selective hybridisation is typically achieved using conditions of medium to high stringency, for example 0.03M sodium chloride and 0.003M sodium citrate at from about 50°C to about 60°C. In a preferred aspect, the present invention covers nucleotide sequences that can hybridise to the nucleotide sequence of the present invention under stringent conditions (such as, for example, 65°C and 0.1 SSC) to the nucleotide sequence presented herein (including complementary sequences of those presented herein).

FRAGMENT

The term "fragment" indicates that the polypeptide comprises a fraction of the wild type amino acid sequence. It may comprise one or more large contiguous sections of sequence or a plurality of small sections. Preferably the polypeptide comprises at least 50%, more preferably at least 65%, most preferably at least 80% of the wild-type sequence.

HETEROLOGOUS PROTEIN/MOLECULE

In contrast to the poor immunogenicity of the A subunit alone, both LTB and CTB are exceptionally potent immunogens. Because of their immunogenicity, both LTB and CTB have been used as carriers for other epitopes and antigens (Nashar *et al* Vaccine 1993;11(2):235-40) and have been used as components of vaccines against cholera and *E. coli* mediated diarrhoeal diseases (Jetborn *et al* 1992 Vaccine 10: 130). The CTB produced by the expression system as described herein may also be used as a carrier for other immunogeneic or tolerogeneic molecules, such as heterologous molecules, which may be coupled to CTB by chemical conjugation or which may be prepared as part of a chimeric protein.

As used herein, the term "heterologous molecule" refers to a molecule which is typically from a different species to the host cell, but may be from a different or unrelated strain of the same species. The host cell may be engineered to express more than one heterologous polypeptide, in which case the polypeptides may be from the same organism or from different organisms. In a preferred embodiment of the

invention, a heterologous nucleotide sequence encodes a heterologous antigen of a pathogen. In another preferred embodiment, two or more heterologous antigens from different pathogens may be expressed. The heterologous DNA or heterologous polypeptide may be a complete protein or a part of a protein containing an epitope. In one embodiment of the invention, the heterologous polypeptide may be the non-toxic component or form of CT or LT. In another embodiment, the heterologous antigen may be an ETEC antigen such as CFA1, CFAII (CS1, CS2, CS3), CFA IV (CS4, CS5, CS6) fimbrial antigen. In yet another embodiment, the heterologous antigen may be expressed or be prepared as part of a fusion protein. In this regard, the fusion protein may involve two or more different antigens or an antigen and a region designed to increase the immunogenicity of a heterologous polypeptide. The heterologous antigen may be selected from the group consisting of viruses, bacteria, fungi, proteins, polypeptides or immunogenic portions thereof. In another embodiment, the immunogenic component is selected from the group consisting of *Bordetella* pertussis toxin subunit S2, S3, S4, S5, Diphtheria toxin fragment B, *E.coli* fimbria K88, K99, 987P, F41, CFA I, CFA II (CS 1, CS2, CS3), CFA IV (CS4, CS5, CS6).

In some embodiments of the invention the heterologous polypeptide encoded by the plasmid to be stabilised may be other than, or in addition to, sequences encoding a heterologous antigen. For example, the polypeptide may regulate or turn on expression of the heterologous antigen encoded by sequences on the bacterial chromosome or a second plasmid. Alternatively, or in addition, the heterologous polypeptide encoded by the plasmid may be a selection marker or a polypeptide required for optimal growth of the bacterium carrying the plasmid. In the case of the heterologous polypeptide playing a regulatory role it may bind to and activate, or increase expression from, the sequences encoding the heterologous antigen. The regulation may be inducible so that expression of the antigen is only activated at an appropriate time, for example when the bacteria are at an appropriate stage of growth or administered to the host to be vaccinated. This may help avoid, or reduce, early selection pressure against bacteria carrying the plasmid until expression is induced.

As indicated above, one or more heterologous molecules may be coupled to the CTB produced by the expression system as described herein by chemical conjugation or which may be prepared as part of a chimeric protein. In one embodiment of the present invention, chemical coupling is carried out using a functional cross-linking reagent, such as a heterobifunctional cross-linking reagent. More preferably the cross-linking agent is N-y (-maleimido-butyroxyl)succinimide ester (GMBS) or N-succinimidyl- (3-pyridyl-dithio)-propionate (SPDP). The term "coupling" includes direct or indirect linkage, for example, by the provision of suitable spacer groups. By way of example, the coupled components may be covalently linked, to form a single active moiety/entity. Alternatively, the coupled components may also be linked to another entity. WO 95/10301 teaches how antigens may be coupled either directly or indirectly to a mucosa-binding molecule.

Method have also been described for making fusion proteins based on CTB or LTB wherein nucleic acids encoding for either or both of T or B epitopes of a heterologous

antigen of interest are genetically fused to coding sequences for either or both of the N-or C-terminus of CTB, or placed in an intrachain position in the CTB or LTB coding sequence, or to analogous positions in CTA or LTA (Backstrom *et al.*, Gene 1995; 165: 163-171, Bäckström *et al.*, Gene 1994; 149: 211-217, Schödel *et al.*, Gene 1991; 99: 255259). Methods have also been described for fusing peptides to the carboxy or amino ends of CTA or LTA and for co-expressing these fusion proteins with CTB or LTB (Sanchez *et al.* FEBS Lett. 1986; 208: 194-198, Sanchez *et al.* FEBS Lett. 1997; 401: 95-97).

By way of example, genetic fusions may be prepared using a vector that contains a promoter for expressing the fusion protein, the DNA sequence of the cholera toxin binding subunit CTB, and an immunogenic peptide coding sequence. The CTB and the immunogenic peptide coding sequence are linked such that they were in the proper reading frame producing a fusion protein. The fusion protein is expressed, secreted, and purified for use as a vaccine. Hybrid CTB/LTB proteins may also be prepared according to the teachings in WO 96/34893 or in accordance with any known method in the art. These expressed hybrid proteins may include a mature CTB sequence in which the amino acid residues are substituted with the corresponding amino acid residues of mature LTB which impart LTB specific epitopes characteristic to said immunogenic mature CTB (resulting in, for example, a hybrid molecule called LCTBA). Conversely, the hybrid protein may include a mature LTB sequence in which the amino acid residues are substituted with the corresponding amino acid residues of mature CTB which impart CTB specific epitopes characteristic to said immunogenic mature LTB (resulting in, for example, a hybrid molecule called LCTBB). In addition a third hybrid protein is envisaged which combines an LCTBA molecule and a LCTBB molecule (see WO 96/34893 and Lebens *et al* (1996) Infect and Immunity 64(6); 2144-2150).

METHOD OF MAKING CTB

Examples of a gene encoding CTB include but are not limited to the CTB gene presented in Figure 16 and those specified under GI Accession No in Table 1. The CTB gene is inserted in an expression vector. The stable expression vector may be made and transformed into bacterium using conventional techniques.

As used herein, the term "transformation", refers to the insertion of an exogenous polynucleotide into a host cell, heterologous gene, nucleotide sequence, such as a DNA or RNA sequence so that the host cell will express the introduced gene or sequence which is typically an RNA coded by the introduced gene or sequence, but also a protein or an enzyme coded by the introduced gene or sequence. Any method may be used for the insertion such as but not limited to direct uptake, transduction, f-mating, use of CaCl2 or other agents, such as divalent cations and DMSO or electroporation. The heterologous or exogeneous polynucleotide may be maintained as a non-integrated vector, for example, a plasmid, or alternatively, may be integrated into the host genome.

Transformation procedures usually vary with the bacterial species to be transformed. See e.g., [Masson *et al.* (1989) FEMS Microbiol. Lett. 60:273; Palva *et al.* (1982) Proc. Natl. Acad. Sci. USA 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT

Publ. No. Wo 84/04541, Bacillus], [Miller *et al.* (1988) Proc. Natl. Acad. Sci. 85:856; Wang *et al.* (1990) J. Bacteriol. 172:949, Campylobacter], [Cohen *et al.* (1973) Proc. Natl. Acad. Sci. 69:2110; Dower *et al.* (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of Escherichia coli with ColE1-derived plasmids. In Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H. W. Boyer and S. Nicosia); Mandel *et al.* (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; Escherichia], [Chassy *et al.* (1987) FEMS Microbiol. Lett. 25 44:173 Lactobacillus]; [Fiedler *et al.* (1988) Anal. Biochem 170:38, Pseudomonas]; [Augustin *et al.* (1990) FEMS Microbiol. Lett. 66:203, Staphylococcus], [Barany *et al.* (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: Streptococcal Genetics (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) Infec. Immun. 32:1295; Powell *et al.* (1988) Appl. Environ. Microbiol. 54:655; Somkuti *et al.* (1987) Proc. 4th Evr. Cong. Biotechnology 1:412, Streptococcus].

A host cell that receives and expresses introduced DNA or RNA has been "transformed" and is a "transformant" or a "clone". The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell or cells of a different genus or species.

Means for introducing the stable expression vector into prokaryotic host cells, such as *V. cholerae* and *E. coli* host cells are known in the art. Examples of suitable methods include but are not limited to electroporation, conjugation and electrophoresis. The transformed colonies may be screened and selected for correct uptake using standard screening and selection procedures. The expression of the CTB is designed so that CTB is overproduced and accumulates in the growth medium.

After culturing, the CTB subunit protein produced by the expression system of the present invention as described herein may be purified by, for example, chromatography, precipitation, and/or density gradient centrifugation. The thus obtained CTB protein may be used as a vaccine or for the production of antibodies directed against said peptides, which can be used for passive immunization.

The CTB produced by the expression system as described herein may be purified from the culture filtrate using standard ammonium sulphate precipitation, ion-exchange and affinity chromatography techniques (as outlined in WO 01/27144). The CTB is characterised using GM-1 ELISA, colorimetric protein assays (A280, Lowry, Bradford, BCA), Western Blots and Single radial immunodiffusion (SRI) and Mancini test (as described in the Examples). Preferably, purified material substantially free of contaminants is at least 50% pure; more preferably, at least 90% pure, and more preferably still at least 99% pure. Purity can be evaluated by chromatography, gel electrophoresis, immunoassay, composition analysis, biological assay, and other methods known in the art.

A significant advantage of producing CTB subunit protein using the expression system as described herein rather than by isolating and purifying a CTB subunit protein from natural sources is that equivalent quantities of the CTB protein can be

produced by using less starting material than would be required for isolating the protein from a natural source. Producing the CTB protein by the expression system as described herein also permits the CTB protein to be isolated in the absence of some molecules, such as antibiotics or antibiotic selection markers normally present in a typical host cells expression system resulting in a safer product for human use. Also, production of the CTB protein from genetically detoxified host cells as described herein provides a safer method of producing CTB because the host cell is unlikely to revert to a toxic form. The expression system for producing CTB as described herein is also advantageous because it results in increased production capacity and reduced costs.

ISOLATED CTB

The isolated stably expressed CTB obtainable by the method as described herein is essentially free of antibiotic residues because the expression system does not express an antibiotic resistance marker and therefore the use of antibiotic additives in the expression system is cessary.

CTB PREPARATION ESSENTIALLY FREE OF ANTIBIOTIC RESIDUES

As used herein, the term "CTB preparation essentially free of antibiotic residues" refers to a CTB preparation obtainable from an expression system which does not express an antibiotic resistance marker and/or which does not require the presence of antibiotics (such as but not limited to ampicillin, streptomycin, sulphmethoxazole, kanamycin, trimetheprim, and tetracycline) to maintain the expression system.

VACCINES

The CTB produced using the expression system of the present invention may be used either alone or in combination with other antigens in the preparation of cholera toxin B based vaccines. Studies with recombinant CTB (rCTB) provide seemingly contradictory data. On the one hand, there is little question that rCTB can act as an immunogen or as a mucosal adjuvant in that it enhances humoural antibody responses to antigen when both are administered by an intranasal route (see Wu *et al* 1998 Vaccine 16: 286 and Tochikubo *et al* 1998 Vaccine 16: 150). On the other hand, there is also good evidence that rCTB may also act as an immunosuppressant or an immunomodulator which is capable of re-directing or down-regulating an immune response pathway (such as the Th1 immune response pathway). By way of example, when rCTB is conjugated to antigen, it enhances oral tolerance induced by the antigen (see Sun *et al* 1994: PNAS (USA) 91: 10795; Sun *et al* 1996 PNAS (USA) 93: 7196; and Bergerot *et al* 1997 PNAS (USA) 94: 4610). In addition, WO 01/27144 teaches that rCTB can be used in the treatment of autoimmunity and the induction of tolerance whilst WO 01/34175 teaches that rCTB can be used in the prevention and/or treatment of Chrohn's disease. Accordingly the CTB produced by the expression system as described herein has uses which include but are not limited to uses as an immunogen, an immunomodulator, a tolerogen, an adjuvant, a carrier for other heterologous antigens, a diagnostic reagent and a receptor blocking agent.

CTB AS AN IMMUNOGEN

The CTB molecule as described herein which may be used to form the vaccines of the present invention may be the whole protein, an immunogenic fragment of the whole

CTB protein, a mutagenized form of the CTB, or a fusion protein comprising the CTB protein or a fragment thereof and a suitable fusion partner. In one embodiment, the CTB may be used in conjunction with inactivated whole *V. cholerae* cells as a combination oral vaccine intended to stimulate at least an anti-bacterial immune response and/or an antitoxic antibody immune response.

CHOLERA TOXIN (CT)

It is well known in the art that the cholera endotoxin (CT) is a 84, 000 molecular weight protein composed of one A subunit and five B subunits. The toxic A subunit induces the enzymatic changes (due to its ADP-ribosylating activity) which lead to fluid secretion and diarrhoea while the non-toxic B subunit is the immunogenic moiety that binds to the GM-1 ganglioside receptor for the toxin on intestinal epithelial cells (Holmgren J Nature (1981) 292; 413). Each B subunit has a high binding affinity for the toxin's cell surface receptor, ganglioside GM-1. It has been shown that purified B subunit given either orally or parenterally stimulates the appearance of serum IgA antitoxin in intestinal fluid. In addition, neutralising antibodies raised against the holotoxin react mainly with the B subunit. The major advantage of using oral vaccines comprising CTB to stimulate antitoxic immunity includes its complete safety as there is no potential for reversion to toxin as exists with toxoids.

ORAL VACCINES

CTB is an effective oral immunizing agent which in a large field trial has been shown to afford protection against both cholera and diarrhoea caused by LT enterotoxigenic *E.coli* (J Clemens *et al.*, Lancet ii:124-127, 1986). The separation of the B subunit from the A subunit excludes any risk of reversion to toxicity, and CTB has been administered orally to more than 25,000 people without any side effects. These features have made CTB an important component, together with killed whole cholera vibrios, of a new oral cholera vaccine.

An oral vaccine comprising the CTB produced by the expression system as described herein is advantageous as the CTB subunit aids in the uptake of the vaccine from the digestive tract across mucosal membranes. By way of example, a dead or inactivated cholera vaccine consisting of whole cells supplemented with the B subunit of cholera toxin is currently available (Holmgren *et al* Current topics in Microbiology and Immunology 1989: 146). This vaccine, which is marketed under the registered name Dukoral® is an oral (liquid) vaccine product comprising a mixture of four inactivated whole cell *V. cholerae* strains and recombinant CTB (rCTB). The *V. cholerae* strains are inactivated so that they do not cause *V. cholerae* associated disease, such as diarrhoea, while nevertheless being capable of inducing an immune response.

The recombinant B subunit (rCTB) of the cholera toxin is an essential component of the Dukoral® product and is the only component of the oral cholera vaccine that is made with recombinant DNA techniques. The vaccine has indications for prevention of diarrhoea induced by enterotoxigenic *E. coli* bacteria (ETEC), especially ETEC bacteria producing heat-labile enterotoxin (LT) to the homology between the CTB enterotoxin and *E. coli* heat labile enterotoxin (LT). The inactivated cholerae strains are: *V. cholerae* 01 Inaba classic biotype (heat inactivated); *V. cholerae* E1 Tor

biotype (formalin inactivated); *V. cholerae* 01 Ogawa biotype (heat inactivated); and *V. cholerae* 01 Ogawa classic biotype (formalin inactivated). The liquid suspension contains a total concentration of 1 x 10^{11} vibrios and 1mg rCTB in a pharmaceutical buffer. This liquid suspension is packaged with a sachet of bicarbonate of soda (to neutralise stomach acidity) and both components are added to water and drunk.

Without wishing to be bound by theory, the Dukoral® vaccine acts by inducing antibodies against both the bacterial components and the CTB component. The anti-bacterial intestinal antibodies prevent the bacteria from attaching to the intestinal wall thereby impeding colonisation of *V. cholerae* 01. The rCTB component of Dukoral® appears to act as a principal antigen to elicit neutralising antibody protection against the cholera enterotoxin. The anti-toxin intestinal antibodies prevent the cholera toxin from binding to the intestinal mucosal surface thereby preventing the toxin-mediated diarrhoeal symptoms. The stimulation of both antitoxin and antibacterial immunity provide synergistic protection. In addition, the antibody response against the rCTB component cross-reacts with LT, thus providing antibody protection against infection by ETEC bacteria.

In one embodiment, the CTB as described herein or a composition comprising CTB as described herein may be used as a vaccine, preferably a mucosal vaccine, to stimulate at least both anti-bacterial and antitoxic antibody production against a enterotoxin induced illness, such as that caused by a cholera toxin or an ETEC toxin.

ETEC VACCINE

An oral ETEC vaccine composition comprising the CTB produced by the expression system as described herein is advantageous as CTB aids in the uptake of the vaccine from the digestive tract across mucosal membranes. By way of example, the ETEC vaccine composition may comprise naturally selected or recombinant *E. coli* host cell strains expressing at least three different types of colonisation factor antigens (CFAs) selected from the group consisting of CFA1, CFA II (CS1, CS2 and CS3) and CFA IV (CS4, CS5 and CS6) together with a defined amount of the CTB which is produced as described herein. The ETEC vaccine composition may be prepared according to the disclosure in WO 00/37106 and/or WO 03/022306.

CTB AS AN ADJUVANT

The CTB produced by the expression system as described herein may be useful as an adjuvant when used in combination with other vaccine components. Typically, an adjuvant is a separate moiety from the antigen, although a single molecule like CTB can have both adjuvant and antigenic properties. It is known that the cholera toxin and its B subunit (CTB) have adjuvant properties when used as either an intramuscular or possibly as an oral immunogen (Elson and Dertzbaugh, J Immunol. 1995; 154 (3): 1032-40; Trach *et al.*, Lancet. 1997; 349 (9047): 231-5).

As used herein, the term "adjuvant" refers to any material or composition capable of specifically or non-specifically altering in ways which include but are not limited to enhancing, directing, redirecting, potentiating or initiating an antigen-specific or non-specific immune response. Without being bound by theory, the inclusion of an adjuvant is advantageous because the adjuvant may help to enhance the cell mediated

immune (CMI) response to a heterologous antigen by directing and/or diverting a Th2 response to a Th1 response and/or directing a specific effector associated mechanisms with the consequent generation and maintenance of an enhanced CMI response (see, for example the teachings in WO 97/02045 and US2001/0036917A1).

The inclusion of an adjuvant, such as CTB with a heterologous antigen in a vaccine composition is also advantageous because it may result in a lower dose or fewer doses of the heterologous antigen being necessary to achieve the desired CMI response in the subject to which the composition is administered, or it may result in a qualitatively and/or quantitatively different immune response in the subject. The effectiveness of a CTB adjuvant can be determined by administering the CTB adjuvant with the heterologous antigen in parallel with the antigen alone to animals and comparing antibody and/or cellular-mediated immunity in the two groups using standard assays such as radioimmunoassay, ELISAs, CTL assays, and the like, which are all well known in the art. US Patent No 5, 182, 109 describes the intranasal administration of a number of vaccine preparations mixed with CTB as an adjuvant to promote an immune response.

In another preferred embodiment, vaccines comprising the CTB as described herein may be used as an non-specific vaccine component (for example, as a vaccine carrier) in other vaccine compositions

CTB AS A PEPTIDE/PROTEIN CARRIER

The CTB produced by the expression system as described herein may also be used as a carrier for other immunogeneic or tolerogeneic molecules (such as, for example, heterologous molecules which may be coupled to CTB by chemical conjugation or as part of a chimeric protein). In some embodiments, the whole CTB molecule as described herein may be used as the carrier component. In other embodiments, a functionally active fragment of the CTB molecule may be used. The present invention also contemplates the use of mutagenized forms of the CTB molecule as carrier components.

Vaccines which couple the CTB protein in some way to an immunogen of choice may be particularly useful for any type of disease which is targeted to mucosal surfaces. Production of a specific vaccine typically involves steps such as but not limited to:

1) identifying an antigen or immunogen; 2) coupling the antigen or immunogen to CTB in some way (chemically, hydrophobically, or genetically) ; 3) isolating the coupled CTB/immunogen and confirming that it still binds to Gm-1 ; 4) identifying that it works *in vivo* by injecting it into a mouse or other model animal ; and 5) testing it for efficacy in humans or other primates. Examples of vaccine targets include but are not limited to those shown in Table 3 of WO 01/27144. The CTB of the present invention can be chemically, hydrophobically, or genetically coupled to a heterologous antigen and fed to animal models to test the efficacy.

CTB AS A TOLEROGEN/IMMUNOMODULATOR

In another preferred embodiment, the CTB as described herein may be used either alone or in combination with one or more antigens to down-regulate or de-sensitise or suppress an allergic/autoimmune response by modulating the immune response.

As used herein, the term "modulating" includes but is not limited to any treatment which may prevent, suppress, alleviate, restore, elevate, direct, re-direct or modify the immune response. By way of example, WO 95/10301 teaches that antigens linked to a mucosa-binding molecule may also induce immunological tolerance to a specific antigen. As used herein, "immunological tolerance" refers to a reduction in immunological reactivity of a host towards a specific antigen or antigens. In addition, WO 01/27144 and WO 01/34175 teach that CTB alone maybe used in the treatment of autoimmunity and the induction of tolerance. WO 02/093998 also demonstrates that CTB can suppress established IgE mediated allergic reactions when administered transcutaneously by entering through the dermis of the skin without perforation of the skin.

COMPOSITIONS

The CTB as described herein may be formulated as a composition. In one embodiment, the composition is a pharmaceutical composition. In another embodiment, the composition is a therapeutic composition. In another preferred embodiment, the composition is an immunotherapeutic composition. In an even more preferred embodiment, the composition is a vaccine composition. The composition can comprise a therapeutically or prophylactically effective amount of a CTB of the as described herein. The composition may also comprise a carrier such as a pharmaceutically or immunologically or therapeutically acceptable carrier. Pharmaceutically acceptable carriers or immunologically or therapeutically acceptable carriers are determined in part by the particular composition being administered as well as by the particular method used to administer the composition.

Pharmaceutically acceptable carriers include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes) and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents (adjuvants).

The composition may, in addition to the immunising component(s), comprise a vehicle, such as physiological saline solution, and other components frequently used in vaccines such as buffers and adjuvants. Useful vehicles, buffers, adjuvants and other components are disclosed in, for example, the European and US Pharmacopoeia. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions or vaccine compositions or immunotherapeutic compositions of the present invention. These compositions may additionally contain one or more adjuvants and/or pharmaceutically acceptable diluents.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to one or more of the following set forth below::

Mineral Containing Compositions

Mineral containing compositions suitable for use as adjuvants in the invention include mineral salts, such as aluminium salts and calcium salts. The invention includes mineral salts such as hydroxides (*e.g.* oxyhydroxides), phosphates (*e.g.* hydroxyphoshpates, orthophosphates), sulphates, *etc.* (*e.g.* see chapters 8 & 9 of *Vaccine design:the subunit and adjuvant approach* (1995) Powell & Newman. ISBN 0-306-44867-X.), or mixtures of different mineral compounds, with the compounds taking any suitable form (*e.g.* gel, crystalline, amorphous, *etc.*), and with adsorption being preferred. The mineral containing compositions may also be formulated as a particle of metal salt. See WO00/23105.

Oil-Emulsions

Oil-emulsion compositions suitable for use as adjuvants in the invention include squalene-water emulsions, such as MF59 (5% Squalene, 0.5% Tween 80, and 0.5% Span 85, formulated into submicron particles using a microfluidizer). See WO90/14837.

Complete Freund's adjuvant (CFA) and incomplete Freund's adjuvant (IFA) may also be used as adjuvants in the invention (see WO 90/14837).

Saponin Formulations

Saponin formulations, may also be used as adjuvants in the invention. Saponins are a heterologous group of sterol glycosides and triterpenoid glycosides that are found in the bark, leaves, stems, roots and even flowers of a wide range of plant species. Saponin from the bark of the *Quillaia saponaria* Molina tree have been widely studied as adjuvants. Saponin can also be commercially obtained from *Smilax ornata* (sarsaprilla), *Gypsophilla paniculata* (brides veil), and *Saponaria officianalis* (soap root). Saponin adjuvant formulations include purified formulations, such as QS21, as well as lipid formulations, such as ISCOMs.

Saponin compositions have been purified using High Performance Thin Layer Chromatography (HP-LC) and Reversed Phase High Performance Liquid Chromatography (RP-HPLC). Specific purified fractions using these techniques have been identified, including QS7, QS17, QS18, QS21, QH-A, QH-B and QH-C. Preferably, the saponin is QS21. A method of production of QS21 is disclosed in U.S. Patent No. 5,057,540. Saponin formulations may also comprise a sterol, such as cholesterol (see WO 96/33739).

Combinations of saponins and cholesterols can be used to form unique particles called Immunostimulating Complexs (ISCOMs). ISCOMs typically also include a phospholipid such as phosphatidylethanolamine or phosphatidylcholine. Any known saponin can be used in ISCOMs. Preferably, the ISCOM includes one or more of Quil A, QHA and QHC. ISCOMs are further described in EP 0 109 942, WO 96/11711 and WO 96/33739. Optionally, the ISCOMS may be devoid of additional detergent. See WO00/07621.

A review of the development of saponin based adjuvants can be found at Barr, et al., "ISCOMs and other saponin based adjuvants", Advanced Drug Delivery Reviews (1998) 32:247 – 271. See also Sjolander, et al., "Uptake and adjuvant activity of

orally delivered saponin and ISCOM vaccines", Advanced Drug Delivery Reviews (1998) 32:321 – 338.

Virosomes and Virus Like Particles (VLPs)

Virosomes and Virus Like Particles (VLPs) can also be used as adjuvants in the invention. These structures generally contain one or more proteins from a virus optionally combined or formulated with a phospholipid. They are generally non-pathogenic, non-replicating and generally do not contain any of the native viral genome. The viral proteins may be recombinantly produced or isolated from whole viruses. These viral proteins suitable for use in virosomes or VLPs include proteins derived from influenza virus (such as HA or NA), Hepatitis B virus (such as core or capsid proteins), Hepatitis E virus, measles virus, Sindbis virus, Rotavirus, Foot-and-Mouth Disease virus, Retrovirus, Norwalk virus, human Papilloma virus, HIV, RNA-phages, Qß-phage (such as coat proteins), GA-phage, fr-phage, AP205 phage, and Ty (such as retrotransposon Ty protein p1). VLPs are discussed further in WO 03/024480, WO 03/024481, and Niikura et al., "Chimeric Recombinant Hepatitis E Virus-Like Particles as an Oral Vaccine Vehicle Presenting Foreign Epitopes", Virology (2002) 293:273 – 280; Lenz et al., "Papillomarivurs-Like Particles Induce Acute Activation of Dendritic Cells", Journal of Immunology (2001) 5246 – 5355; Pinto, et al., "Cellular Immune Responses to Human Papillomavirus (HPV)-16 L1 Healthy Volunteers Immunized with Recombinant HPV-16 L1 Virus-Like Particles", Journal of Infectious Diseases (2003) 188:327 – 338; and Gerber et al., "Human Papillomavrisu Virus-Like Particles Are Efficient Oral Immunogens when Coadministered with Escherichia coli Heat-Labile Entertoxin Mutant R192G or CpG", Journal of Virology (2001) 75(10):4752 – 4760. Virosomes are discussed further in, for example, Gluck et al., "New Technology Platforms in the Development of Vaccines for the Future", Vaccine (2002) 20:B10 –B16.

Bacterial or Microbial Derivatives

Adjuvants suitable for use in the invention include bacterial or microbial derivatives such as:

(1) *Non-toxic derivatives of enterobacterial lipopolysaccharide (LPS)*

Such derivatives include Monophosphoryl lipid A (MPL) and 3-O-deacylated MPL (3dMPL). 3dMPL is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains. A preferred "small particle" form of 3 De-O-acylated monophosphoryl lipid A is disclosed in EP 0 689 454. Such "small particles" of 3dMPL are small enough to be sterile filtered through a 0.22 micron membrane (see EP 0 689 454). Other non-toxic LPS derivatives include monophosphoryl lipid A mimics, such as aminoalkyl glucosaminide phosphate derivatives *e.g.* RC-529. See Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278.

(2) *Lipid A Derivatives*

Lipid A derivatives include derivatives of lipid A from *Escherichia coli* such as OM-174. OM-174 is described for example in Meraldi et al., "OM-174, a New Adjuvant with a Potential for Human Use, Induces a Protective Response with Administered with the Synthetic C-Terminal Fragment 242-310 from the circumsporozoite protein

of Plasmodium berghei", Vaccine (2003) 21:2485 – 2491; and Pajak, et al., "The Adjuvant OM-174 induces both the migration and maturation of murine dendritic cells in vivo", Vaccine (2003) 21:836 – 842.

(3) *Immunostimulatory oligonucleotides*

Immunostimulatory oligonucleotides suitable for use as adjuvants in the invention include nucleotide sequences containing a CpG motif (a sequence containing an unmethylated cytosine followed by guanosine and linked by a phosphate bond). Bacterial double stranded RNA or oligonucleotides containing palindromic or poly(dG) sequences have also been shown to be immunostimulatory.

The CpG's can include nucleotide modifications/analogs such as phosphorothioate modifications and can be double-stranded or single-stranded. Optionally, the guanosine may be replaced with an analog such as 2'-deoxy-7-deazaguanosine. See Kandimalla, et al., "Divergent synthetic nucleotide motif recognition pattern: design and development of potent immunomodulatory oligodeoxyribonucleotide agents with distinct cytokine induction profiles", Nucleic Acids Research (2003) 31(9): 2393 – 2400; WO 02/26757 and WO 99/62923 for examples of possible analog substitutions. The adjuvant effect of CpG oligonucleotides is further discussed in Krieg, "CpG motifs: the active ingredient in bacterial extracts?", Nature Medicine (2003) 9(7): 831 – 835; McCluskie, et al., "Parenteral and mucosal prime-boost immunization strategies in mice with hepatitis B surface antigen and CpG DNA", FEMS Immunology and Medical Microbiology (2002) 32:179 – 185; WO 98/40100; U.S. Patent No. 6,207,646; U.S. Patent No. 6,239,116 and U.S. Patent No. 6,429,199.

The CpG sequence may be directed to TLR9, such as the motif GTCGTT or TTCGTT. See Kandimalla, et al., "Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic CpG DNAs", Biochemical Society Transactions (2003) 31 (part 3): 654 – 658. The CpG sequence may be specific for inducing a Th1 immune response, such as a CpG-A ODN, or it may be more specific for inducing a B cell response, such a CpG-B ODN. CpG-A and CpG-B ODNs are discussed in Blackwell, et al., "CpG-A-Induced Monocyte IFN-gamma-Inducible Protein-10 Production is Regulated by Plasmacytoid Dendritic Cell Derived IFN-alpha", J. Immunol. (2003) 170(8):4061 – 4068; Krieg, "From A to Z on CpG", TRENDS in Immunology (2002) 23(2): 64 – 65 and WO 01/95935. Preferably, the CpG is a CpG-A ODN.

Preferably, the CpG oligonucleotide is constructed so that the 5' end is accessible for receptor recognition. Optionally, two CpG oligonucleotide sequences may be attached at their 3' ends to form "immunomers". See, for example, Kandimalla, et al., "Secondary structures in CpG oligonucleotides affect immunostimulatory activity", BBRC (2003) 306:948 – 953; Kandimalla, et al., "Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic GpG DNAs", Biochemical Society Transactions (2003) 31(part 3):664 – 658; Bhagat et al., "CpG penta- and hexadeoxyribonucleotides as potent immunomodulatory agents" BBRC (2003) 300:853 – 861 and WO 03/035836.

(4) *ADP-ribosylating toxins and detoxified derivatives thereof.*

Bacterial ADP-ribosylating toxins and detoxified derivatives thereof may be used as adjuvants in the invention. Preferably, the protein is derived from *E. coli* (i.e., E. coli heat labile enterotoxin "LT), cholera ("CT"), or pertussis ("PT"). The use of

detoxified ADP-ribosylating toxins as mucosal adjuvants is described in WO 95/17211 and as parenteral adjuvants in WO 98/42375. Preferably, the adjuvant is a detoxified LT mutant such as LT-K63, LT-R72, and LTR192G. The use of ADP-ribosylating toxins and detoxified derivaties thereof, particularly LT-K63 and LT-R72, as adjuvants can be found in the following references, each of which is specifically incorporated by reference herein in their entirety: Beignon, et al., "The LTR72 Mutant of Heat-Labile Enterotoxin of Escherichia coli Enahnces the Ability of Peptide Antigens to Elicit CD4+ T Cells and Secrete Gamma Interferon after Coapplication onto Bare Skin", Infection and Immunity (2002) 70(6):3012 – 3019; Pizza, et al., "Mucosal vaccines: non toxic derivatives of LT and CT as mucosal adjuvants", Vaccine (2001) 19:2534 – 2541; Pizza, et al., "LTK63 and LTR72, two mucosal adjuvants ready for clinical trials" Int. J. Med. Microbiol (2000) 290(4-5):455-461; Scharton-Kersten et al., "Transcutaneous Immunization with Bacterial ADP-Ribosylating Exotoxins, Subunits and Unrelated Adjuvants", Infection and Immunity (2000) 68(9):5306 – 5313; Ryan et al., "Mutants of Escherichia coli Heat-Labile Toxin Act as Effective Mucosal Adjuvants for Nasal Delivery of an Acellular Pertussis Vaccine: Differential Effects of the Nontoxic AB Complex and Enzyme Activity on Th1 and Th2 Cells" Infection and Immunity (1999) 67(12):6270 – 6280; Partidos et al., "Heat-labile enterotoxin of Escherichia coli and its site-directed mutant LTK63 enhance the proliferative and cytotoxic T-cell responses to intranasally co-immunized synthetic peptides", Immunol. Lett. (1999) 67(3):209 – 216; Peppoloni et al., "Mutants of the Escherichia coli heat-labile enterotoxin as safe and strong adjuvants for intranasal delivery of vaccines", Vaccines (2003) 2(2):285 – 293; and Pine et al., (2002) "Intranasal immunization with influenza vaccine and a detoxified mutant of heat labile enterotoxin from Escherichia coli (LTK63)" J. Control Release (2002) 85(1-3):263 – 270. Numerical reference for amino acid substitutions is preferably based on the alignments of the A and B subunits of ADP-ribosylating toxins set forth in Domenighini et al., Mol. Microbiol (1995) 15(6):1165 – 1167, specifically incorporated herein by reference in its entirety.

Human Immunomodulators

Human immunomodulators suitable for use as adjuvants in the invention include cytokines, such as interleukins (*e.g.* IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, *etc.*), interferons (*e.g.* interferon-γ), macrophage colony stimulating factor, and tumor necrosis factor.

Bioadhesives and Mucoadhesives

Bioadhesives and mucoadhesives may also be used as adjuvants in the invention. Suitable bioadhesives include esterified hyaluronic acid microspheres (Singh *et al.* (2001) *J. Cont. Rele.* 70:267-276) or mucoadhesives such as cross-linked derivatives of poly(acrylic acid), polyvinyl alcohol, polyvinyl pyrollidone, polysaccharides and carboxymethylcellulose. Chitosan and derivatives thereof may also be used as adjuvants in the invention. E.g., WO99/27960.

Microparticles

Microparticles may also be used as adjuvants in the invention. Microparticles (*i.e.* a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, and most preferably ~500nm to ~10μm in diameter) formed from materials

that are biodegradable and non-toxic (*e.g.* a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone, *etc.*), with poly(lactide-co-glycolide) are preferred, optionally treated to have a negatively-charged surface (*e.g.* with SDS) or a positively-charged surface (*e.g.* with a cationic detergent, such as CTAB).

Liposomes

Examples of liposome formulations suitable for use as adjuvants are described in U.S. Patent No. 6,090,406, U.S. Patent No. 5,916,588, and EP 0 626 169.

Polyoxyethylene ether and Polyoxyethylene Ester Formulations

Adjuvants suitable for use in the invention include polyoxyethylene ethers and polyoxyethylene esters. WO99/52549. Such formulations further include polyoxyethylene sorbitan ester surfactants in combination with an octoxynol (WO01/21207) as well as polyoxyethylene alkyl ethers or ester surfactants in combination with at least one additional non-ionic surfactant such as an octoxynol (WO01/21152).

Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether (laureth 9), polyoxyethylene-9-steoryl ether, polyoxytheylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether.

Polyphosphazene (PCPP)

PCPP formulations are described, for example, in Andrianov et al., "Preparation of hydrogel microspheres by coacervation of aqueous polyphophazene solutions", Biomaterials (1998) 19(1 – 3):109 – 115 and Payne et al., "Protein Release from Polyphosphazene Matrices", Adv. Drug. Delivery Review (1998) 31(3):185 – 196.

Muramyl peptides

Examples of muramyl peptides suitable for use as adjuvants in the invention include N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), and N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine MTP-PE).

Imidazoquinolone Compounds.

Examples of imidazoquinolone compounds suitable for use adjuvants in the invention include Imiquamod and its homologues, described further in Stanley, "Imiquimod and the imidazoquinolones: mechanism of action and therapeutic potential" Clin Exp Dermatol (2002) 27(7):571 – 577 and Jones, "Resiquimod 3M", Curr Opin Investig Drugs (2003) 4(2):214 – 218.

The invention may also comprise combinations of aspects of one or more of the adjuvants identified above. For example, the following adjuvant compositions may be used in the invention:

(1) a saponin and an oil-in-water emulsion (WO99/11241);

(2) a saponin (e.g.., QS21) + a non-toxic LPS derivative (e.g., 3dMPL) (see WO 94/00153);

(3) a saponin (e.g.., QS21) + a non-toxic LPS derivative (e.g., 3dMPL) + a cholesterol;

(4) a saponin (*e.g.* QS21) + 3dMPL + IL-12 (optionally + a sterol) (WO98/57659);

combinations of 3dMPL with, for example, QS21 and/or oil-in-water emulsions (See European patent applications 0835318, 0735898 and 0761231);

(5) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-block polymer L121, and thr-MDP, either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion.

(6) Ribi™ adjuvant system (RAS), (Ribi Immunochem) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); and

(7) one or more mineral salts (such as an aluminum salt) + a non-toxic derivative of LPS (such as 3dPML).

Aluminium salts and MF59 are preferred adjuvants for parenteral immunisation. Mutant bacterial toxins are preferred mucosal adjuvants.

The immunogenic compositions (e.g. the antigen, pharmaceutically acceptable carrier and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

Typically, the immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect as discussed above under pharmaceutically acceptable carriers.

The CTB produced by the expression system as described herein may be formulated into a pharmaceutical composition or an immunotherapeutic composition or a therapeutic composition or a vaccine composition. Such formulations comprise the CTB combined with a pharmaceutically acceptable carrier, such as sterile water or sterile isotonic saline. Such formulations may be prepared, packaged, or sold in a form suitable for bolus administration or for continuous administration. Injectable formulations may be prepared, packaged, or sold in unit dosage form, such as in ampoules or in multi-dose containers containing a preservative. Formulations include, but are not limited to, suspensions, solutions, emulsions in oily or aqueous vehicles, pastes, and implantable sustained-release or biodegradable formulations. Such formulations may further comprise one or more additional ingredients including, but not limited to, suspending, stabilizing, or dispersing agents.

In one embodiment of a formulation for parenteral administration, the active ingredient is provided in dry (such as, for example, a powder or granules) form for reconstitution with a suitable vehicle (such as, for example, sterile pyrogen-free water) prior to parenteral administration of the reconstituted composition. The pharmaceutical compositions may be prepared, packaged, or sold in the form of a sterile injectable aqueous or oily suspension or solution. This suspension or solution may be formulated according to the known art, and may comprise, in addition to the active ingredient, additional ingredients such as the dispersing agents, wetting agents, or suspending agents described herein. Such sterile injectable formulations may be prepared using a non-toxic parenterally-acceptable diluent or solvent, such as water or 1,3-butane diol, for example. Other acceptable diluents and solvents include, but are not limited to, Ringer's solution, isotonic sodium chloride solution, and fixed oils such as synthetic mono-or di-glycerides. Other parentally-administrable formulations which are useful include those which comprise the active ingredient in microcrystalline form, in a liposomal preparation, or as a component of a biodegradable polymer systems.

Compositions for sustained release or implantation may comprise pharmaceutically acceptable polymeric or hydrophobic materials such as an emulsion, an ion exchange resin, a sparingly soluble polymer, or a sparingly soluble salt.

FORMULATIONS

The CTB produced by the expression system as described herein or variants homologues, derivatives or fragments thereof, are useful alone or as component parts of a conventional vaccine mixtures. A vaccine composition comprising the produced CTB may be formulated using known techniques for formulating vaccine compositions. In one preferred embodiment, the vaccine is advantageously presented for mucosal administration, preferably oral administration, for example as a dried stabilised powder for reconstitution in a suitable buffer prior to administration. The vaccine composition may also contain attenuated bacterial cells. Reconstitution is advantageously effected in a buffer at a suitable pH to ensure the viability of the bacterial component. In order to protect the attenuated bacteria and the vaccine from gastric acidity, a sodium bicarbonate preparation is advantageously administered with each administration of the vaccine. Alternatively the vaccine is presented in a lyophilised encapsulated form. A freeze dried (or lyophilised) preparation is advantageous because it significantly enhances the shelf storage and distribution profile of the product. Procedures for lyophilisation of vaccine compositions are disclosed in WO 03/053463.

In one embodiment, a vaccine composition comprising the CTB produced by the expression system as described herein is in the form of a solid powder. A solid dose formulation is advantageous because it can be stored at room temperature and does not need a continuous cold-chain when distributed in developing countries. A solid dose formulation is also lighter and has a smaller volume and is thus easier to store and to transport. Other advantages of a solid dose formulation include but are not limited to: improved stability, simplified storage and transport, simplified customer handling, lower cost of production, elimination of aseptic production technique and improved capacity of final dosage form. The solid dose formulation may be produced

by methods which include but are not limited to spray drying, wet granulation and lyophilisation. Both spray-drying and wet-granulation methods are advantageous because they have a lower cost than freeze-drying. In this regard, spray-dried/wet granulated powder formulations are cheaper because sterile conditions are not required, there is no requirement for aseptic filling and the formulated products may be stored and transported at room temperature. In addition, after mixture of the powder and the appropriate buffer, there is only one filling operation.

In one embodiment, the formulation comprises the CTB produced by the expression system as described herein, the four inactivated cholera strains as described for the Dukoral® composition and maltodextrin or sucrose. The inclusion of a maltodextrin or sucrose component may help in the stabilisation of the vaccine antigen. The final bulk is spray-dried and blended with the bicarbonate buffer powder or wet-granulated together with the bicarbonate buffer powder yielding a powder with suitable flow characteristics. The powder is readily soluble in water. The powder may be packed in an airtight sachet, suitable for use in areas with hight humidity. Before use, the contents of the sachet will be dissolved in a glass of water. The composition may be stored at room temperature (no cold-chain needed).

In another embodiment, a vaccine composition comprising the CTB as described herein is in the form of a effervescent tablet or an enteric coated tablet.

KITS

Also included in the invention is a kit comprising the CTB as described herein. The kit may also include an adjuvant, which is administered with or as part of the CTB and instructions for administering the CTB. Other preferred components of the kit include an applicator for administering the CTB. As used herein, the term "applicator" refers to any device including but not limited to a hypodermic syringe, gene gun, particle acceleration device, nebulizer, dropper, bronchoscope, suppository, impregnated or coated vaginally-insertable material such as a tampon, douche preparation, solution for vaginal irrigation, retention enema preparation, suppository, or solution for rectal or colonic irrigation for applying the CTB either systemically or mucosally or transdermally or otherwise to the host subject.

The CTB of the present invention may be administered, either alone or as part of a composition, via a variety of different routes. In one embodiment, the CTB is administered alone. In another embodiment, the CTB is administered in admixture with a suitable antigen/tolerogen against which it is desired to alter an immune response by enhancing, directing, re-directing, potentiating, initiation, immomodulating, down-regulating or de-sensitising an immune response. If the CTB and the antigen are not in admixture, it is preferred that they be administered within a short time of each other, at the same site of administration. The composition may be administered as a single dose, a multiple dose, a prime and boost dose or in repeated prime and boost doses.

Certain routes may be favoured for certain compositions, as resulting in the generation of a more enhanced immune response, or as being less likely to induce side effects, or as being easier for administration. The route of administration for a vaccine

composition may vary depending upon the identity of the pathogen or infection to be prevented or treated. Examples of routes of administration include but are not limited to systemic, mucosal, transcutaneous, and transdermal routes of administration.

The CTB of the present invention may be administered via a systemic route or a mucosal route or a transdermal route or it may be administrered directly into a specific tissue such as the liver, bone marrow or into the tumour in the case of cancer therapy. As used herein, the term "systemic administration" includes but is not limited to any parenteral routes of administration. In particular, parenteral administration includes but is not limited to subcutaneous, intraperitoneal, intravenous, intraarterial, intramuscular, or intrasternal injection, intravenous, intraarterial, or kidney dialytic infusion techniques. Preferably, the systemic, parenteral administration is intramuscular injection.

As used herein, the term "mucosal administration" includes but is not limited to oral, intranasal, intravaginal, intrarectal, intratracheal, intestinal and ophthalmic administration. Mucosal routes, particularly intranasal, intratracheal, and ophthalmic are preferred for protection against natural exposure to environmental pathogens such as RSV, flu virus and cold viruses or to allergens such as grass and ragweed pollens and house dust mites. The enhancement of the cell mediated immune (CMI) response will enhance the protective effect against a subsequently encountered target antigen such as an allergen or microbial agent.

In one preferred embodiment of the method, the CTB is administered via a transdermal route. While it is believed that any accepted mode and route of immunization can be employed and nevertheless achieve some advantages in accordance herewith, the examples below demonstrate particular advantages with transdermal administration. In this regard, and without being bound by theory, it is believed that transdermal administration is preferred because it may more efficiently activate the cell-mediated immune (CMI) arm of the immune system by altering, initiating, enhancing or potentiating the immune response. Alternatively, transdermal administration of the CTB either alone or in combination with other elements may more efficiently activate the cell mediated immunomodulatory arm of the immune system which may re-direct the immune response or induce immunologoical tolerance or down-regulate an unwanted immune response in a host mammalian subject which may be associated with an autoimmune disorder, an allergic disorder, a tissue or cell graft rejection event and/or an acute or chronic inflammatory reaction or disorder.

The term "transdermal delivery" intends intradermal (such as, for example, into the dermis or epidermis), transdermal (such as, for example, percutaneous and transmucosal administration, i.e., delivery by passage of an agent into or through skin or mucosal tissue. *See, e.g., Transdermal Drug Delivery: Developmental Issues and Research Initiatives*, Hadgraft and Guy (eds.), Marcel Dekker, Inc., (1989); *Controlled Drug Delivery: Fundamentals and Applications*, Robinson and Lee (eds.), Marcel Dekker Inc.,(1987); and *Transdermal Delivery of Drugs*, Vols. 1-3, Kydonieus and Berner (eds.), CRC Press, (1987). Thus, the term encompasses delivery of an agent using a particle delivery device (e.g., a needleless syringe) such as those described in U.S. Patent No. 5,630,796. In one embodiment, the CTB as described

herein or a composition comprising the CTB, such as a vaccine composition comprising the four inactivated cholera strains as described for the as described for the Dukoral® composition, is administered transdermally.

In another embodiment, the CTB as described herein or a composition comprising the CTB as described herein is administered transcutaneously. As used herein, the term "transcutaneous" is distinguished from conventional topical administration techniques, such as mucosal or transdermal administration because the former requires a mucous membrane (eg lung, mouth, nose, rectum) not found in the skin while the latter requires perforation of the skin through the dermis (see WO 02/064162). In another preferred embodiment of the present invention, the CTB may be administered to cells which have been isolated from the host subject. In this preferred embodiment, preferably the CTB is administered to professional antigen presenting cells (APCs). As used herein, the term "antigen presenting cell" refers to any cell which is an MHC class I bearing cell. Examples of APCs include but are not limited to hematopoietic stem cells, lymphocytes, vascular endothelial cells, respiratory epithelial cells, keratinocytes, skeletal and cardiac muscle cells, neurons, cancer cells respiratory airway epithelial cells, hepatocytes, muscle cells, cardiac myocytes, synoviocytes, primary mammary epithelial cells and post-mitotically terminally differentiated nonreplicating cells such as macrophages or neurons and professional antigen presenting cells (APC) such as dendritic cells or macrophages.

PARTICLE ADMINISTRATION

Particle-mediated methods for delivering protein or nucleic acid preparations are known in the art. Thus, once prepared and suitably purified, the above-described CTB can be coated onto core carrier particles using a variety of techniques known in the art. Carrier particles are selected from materials which have a suitable density in the range of particle sizes typically used for intracellular delivery from a gene gun device. The optimum carrier particle size will, of course, depend on the diameter of the target cells.

The term "core carrier" refers to a carrier on which a guest nucleotide sequences (such as, for example, DNA, RNA) or a protein is coated in order to impart a defined particle size as well as a sufficiently high density to achieve the momentum required for cell membrane penetration, such that the guest molecule can be delivered using particle-mediated techniques (see, e.g., U.S. Patent No. 5,100,792). Core carriers typically include materials such as tungsten, gold, platinum, ferrite, polystyrene and latex. See e.g., *Particle Bombardment Technology for Gene Transfer*, (1994) Yang, N. ed., Oxford University Press, New York, NY pages 10-11. Tungsten and gold particles are preferred. Tungsten particles are readily available in average sizes of 0.5 to 2.0 microns in diameter. Gold particles or microcrystalline gold (e. g., gold powder A1570, available from Engelhard Corp., East Newark, NJ) will also find use with the present invention. Gold particles provide uniformity in size (available from Alpha Chemicals in particle sizes of 1-3 microns, or available from Degussa, South Plainfield, NJ in a range of particle sizes including 0.95 microns). Microcrystalline gold provides a diverse particle size distribution, typically in the range of 0.5-5 microns. However, the irregular surface area of microcrystalline gold provides for highly efficient coating with nucleic acids. A number of methods are known and have

been described for coating or precipitating nucleotide sequences or proteins onto gold or tungsten particles. Most such methods generally combine a predetermined amount of gold or tungsten with, for example, plasmid DNA, CaCl2 and spermidine. The resulting solution is vortexed continually during the coating procedure to ensure uniformity of the reaction mixture. After precipitation of the nucleotide sequence or the protein, the coated particles can be transferred to suitable membranes and allowed to dry prior to use, coated onto surfaces of a sample module or cassette, or loaded into a delivery cassette for use in particular gene gun instruments.

The term "particle delivery device" refers to an instrument which delivers a particulate composition transdermally without the aid of a conventional needle to pierce the skin. Various particle acceleration devices suitable for particle-mediated delivery are known in the art, and are all suited for use in the practice of the invention. Current device designs employ an explosive, electric or gaseous discharge to propel the coated carrier particles toward target cells. The coated carrier particles can themselves be releasably attached to a movable carrier sheet, or removably attached to a surface along which a gas stream passes, lifting the particles from the surface and accelerating them toward the target. An example of a gaseous discharge device is described in U. S. Patent No. 5,204,253. An explosive-type device is described in U. S. Patent No. 4,945,050. One example of a helium discharge-type particle acceleration apparatus is the PowderJect XR instrument (PowderJect Vaccines, Inc., Madison), WI, which instrument is described in U. S. Patent No. 5,120,657. An electric discharge apparatus suitable for use herein is described in U. S. Patent No. 5,149,655. The disclosure of all of these patents is incorporated herein by reference. Alternatively, particulate nucleotide compositions or protein/peptide compositions can administered transdermally using a needleless syringe device. For example, a particulate composition comprising the CTB of the present invention can be obtained using general pharmaceutical methods such as simple evaporation (crystallization), vacuum drying, spray drying or lyophilization.

If desired, the particles can be further densified using the techniques described in commonly owned International Publication No. WO 97/48485, incorporated herein by reference. These particulate compositions can then be delivered from a needleless syringe system such as those described in commonly owned International Publication Nos. WO 94/24263, WO 96/04947, WO 96/12513, and WO 96/20022, all of which are incorporated herein by reference. Delivery of particles comprising antigens or allergens from the above referenced needleless syringe systems is practiced with particles having an approximate size generally ranging from 0.1 to 250 microns, preferably ranging from about 10-70 microns. Particles larger than about 250 microns can also be delivered from the devices, with the upper limitation being the point at which the size of the particles would cause untoward damage to the skin cells. The actual distance which the delivered particles will penetrate a target surface depends upon particle size (e. g., the nominal particle diameter assuming a roughly spherical particle geometry), particle density, the initial velocity at which the particle impacts the surface, and the density and kinematic viscosity of the targeted skin tissue. In this regard, optimal particle densities for use in needleless injection generally range between about 0.1 and 25 g/cm3, preferably between about 0.9 and 1.5 g/cm3, and injection velocities generally range between about 100 and 3,000 m/sec. With appropriate gas pressure, particles having an average diameter of 10-70 Rm can be

accelerated through the nozzle at velocities approaching the supersonic speeds of a driving gas flow.

The particle compositions or coated particles are administered to the individual in a manner compatible with the dosage formulation, and in an amount that will be effective for the purposes of the invention. The amount of the composition to be delivered (e. g., about 0.1 mg to 1 mg, more preferably 1 to 50 ug of the antigen or allergen, depends on the individual to be tested. The exact amount necessary will vary depending on the age and general condition of the individual to be treated, and an appropriate effective amount can be readily determined by one of skill in the art upon reading the instant specification.

Gold or tungsten microparticles can also be used as transporting agents, as described in WO 93/17706, and Tang petal., Nature (1992)356:152. In this particular case, the nucleotide sequence or the protein (such as the CTB subunit protein) is precipitated on the microparticles in the presence of calcium chloride and spermidine, and then the whole is administered by a high-speed jet into the dermis or into the epidermis using an apparatus with no needle, such as those described in U.S. Patent Nos. 4,945,050 and 5,015,580, and WO 94/24243. The quantity of the nucleotide sequence or the protein sequence that can be used to vaccinate a host subject depends on a number of factors such as, for example, the strength of the promoter used to express the antigen, the immunogenicity of the product expressed, the condition of the mammal for whom the administration is intended (e.g., the weight, age, and general state of health), the mode of administration, and the type of formulation. In general, an appropriate dose for prophylactic or therapeutic use in an adult of the human species is from about 1 pg to about 5 mg, preferably from about 10 pg to about 1 mg, most preferably from about 25 pg to about 500 pg. Particle-mediated delivery techniques have been compared to other types of modes of administration, and found markedly superior (Fynan *et al.* (1995) Int. J. Immunopharmacology 17:79-83, Fynan *et al.* (1993) Proc. Natl. Acad. Sci. USA 90:11478-11482, and Raz *et al.* (1994) Proc. Natl. Acad. Sci. USA 91:9519-9523). Such studies have investigated particle-mediated delivery of nucleic acid-based vaccines to both superficial skin and muscle tissue. One possible reason for the markedly better results achieved with the gene gun is that the nucleotide sequence is delivered intracellularly as opposed to the extracellular delivery by intramuscular injection.

HOST MAMMALIAN SUBJECT

The CTB as described herein may be used in the vaccination of a mammalian host, particularly a human host. As used herein, the term "host mammalian subject" refers to any member of the subphylum cordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The terms do not denote a particular age. Thus, both adult and newborn individuals are intended to be covered. The methods described herein are intended for use in any of the above vertebrate species, since the immune systems of all of these vertebrates operate

similarly. If a mammal, the subject will preferably be a human, but may also be a domestic livestock, laboratory subject or pet animal.

PREVENTION AND/OR TREATMENT

The CTB produced by the expression system as described herein may be used either alone or in combination with other components for prophylactic and therapeutic treatment of cholera, enterotoxigenic *E.coli* (ETEC) caused diarrahoea and other diseases, such as but not limited to infectious disease, autoimmune disease and unwanted immune responses associated with autoimmune disorder, an allergic disorder, a tissue or cell graft rejection event and/or an acute or chronic inflammatory reaction or disorder. By way of example, the CTB may be used in prophylactic and/or therapeutic vaccines (including immunotherapeutic vaccines). Although diseases caused by microorganisms which invade via mucosal membrane surfaces are the best targets for CTB vaccines, the CTB produced by the present invention could be used for almost any type of vaccine. Examples of diseases and antigens that could be coupled to CTB for production of a vaccine include but are not limited to those listed in Table 3 of WO 01/27144. It is to be appreciated that all references herein to treatment include curative, palliative and prophylactic treatment.

The CTB described herein may be employed alone as part of a composition, such as but not limited to a pharmaceutical composition or a vaccine composition or a therapeutic composition or an immunotherapeutic composition to prevent and/or treat a T cell mediated immune disorder. The administration of the CTB of the present invention or a composition comprising the CTB may be for either a "prophylactic" or a "therapeutic" purpose. As used herein, the term "treatment" includes any of following: the prevention of infection or reinfection; the reduction or elimination of symptoms; and the reduction or complete elimination of a pathogen. Treatment may be effected prophylactically (prior to infection) or therapeutically (following infection).

Prophylaxis or therapy includes but is not limited to eliciting an effective immune response such as a CMI immune response, and/or alleviating, reducing, curing or at least partially arresting symptoms and/or complications resulting from a T cell mediated immune disorder. When provided prophylactically, the CTB or a composition comprising the CTB of the present invention is typically provided in advance of any symptom. The prophylactic administration of the CTB or a composition comprising the CTB is to prevent or ameliorate any subsequent infection or disease. When provided therapeutically, the CTB or a composition comprising the CTB of the present invention is typically provided at (or shortly after) the onset of a symptom of infection or disease. Thus the CTB or the composition comprising the CTB of the present invention may be provided either prior to the anticipated exposure to a disease causing agent or disease state or after the initiation of an infection or disease.

Whether prophylactic or therapeutic administration (either CTB alone or as part of a composition) is the more appropriate will usually depend upon the nature of the disease or disorder. By way of example, an immunotherapeutic composition comprising the CTB of the present invention could be used in immunotherapy

protocols to actively inducing tumour immunity by vaccination with a tumour cell or its antigenic components. This latter form of treatment is advantageous because the immunity is prolonged and because there is a general belief that one of the best ways to eliminate tumours would be to induce a strong specific antitumour cytotoxic T lymphocyte (CTL) response. On the other hand a vaccine composition will preferably, though not necessarily be used prophylactically to induce an effective CMI response against subsequently encountered antigens or portions thereof (such as epitopes) related to the target antigen.

The dose of CTB as described herein or a composition comprising the CTB administrated to a host subject, in the context of the present invention, should be sufficient to effect a beneficial prophylactic or therapeutic immune response in the subject over time. As used herein, the term a "prophylactically or therapeutically effective dose" means a dose in an amount sufficient to elicit a beneficial response to one or more epitopes of a specific target antigens and/or to alleviate, reduce, cure or at least partially arrest symptoms and/or complications from a T cell mediated immune disorder

DOSAGE

Prophylaxis or therapy can be accomplished by a single direct administration at a single time point or multiple time points. Administration can also be delivered to a single or to multiple sites. Some routes of administration, such as mucosal administration, via ophthalmic drops may require a higher dose. Those skilled in the art can adjust the dosage and concentration to suit the particular route of delivery.

The doses required will differ depending on circumstances such as body weight, interferences with other administered medicaments etc. Generally, the upper limit is not critical unless the dose shows toxicity. By way of example, an infection caused by a microorganism, especially a pathogen, may be prevented by administering an effective dose of a CTB vaccine composition prepared according to the invention. The dosage employed may ultimately be at the discretion of the physician, but will be dependent on various factors including the size and weight of the host and the formulation of the vaccine.

T CELL MEDIATED IMMUNE DISORDERS

The present invention provides a CTB subunit protein as described herein when used alone or compositions comprising the CTB which are useful for preventing and/or treating T cell mediated immune disorders and/or infectious disease caused by a pathogenic agent such as, but not limited to a viral, bacterial, parasitic or other infectious agent and/or a target antigen, such as a tumour associated antigen (TAA) associated with cancer and/or an allergen associated with an allergic disorder. Examples of allergens include, but are not limited to, plant pollens, dust mite proteins, animal dander, saliva and fungal spores. Examples of tumour-associated antigens (TAAs) include, but are not limited to, live or irradiated tumor cells, tumor cell extracts and protein subunits of tumor antigens. The antigen can also be a sperm protein for use in contraception. In some embodiments, the antigen is an environmental antigen. Examples of environmental antigens include, but are not limited to, respiratory syncytial virus ("RSV"), flu viruses and cold viruses. Pathogens

which invade via the mucosa also include those that cause respiratory syncytial virus, flu, other upper respiratory conditions, as well as agents which cause intestinal infections.

Amongst several known examples of other diseases against which an enhanced immune response, in particular, an enhanced CMI response, is important are the following: infection and disease caused by viruses such as but not limited to HIV, herpes simplex, herpes zoster, hepatitis C, hepatitis B, influenza, Epstein-Barr virus, measles, dengue and HTLV-1; diseases caused by bacteria such as but not limited to Mycobacterium tuberculosis and Listeria sp, Chlamydia, Mycobacteria, Plasmodium Falciparum, Legioniella and enteropathogenic, enterotoxigenic, enteroinvasive, enterohaemorrhagic and enteroaggregative *E.coli* and and diseases caused by Pathogenic protozoans which include but which are not limited to malaria, Babesia, Schistosoma, Toxiplasma and Toxocara canis or by the protozoan parasites Toxoplasma and Trypanosoma. Furthermore, the administration regime described herein is expected to be of value in immunising against forms of cancer where CTL cell responses plays a protective role. Examples of cancers of mammals which may be treated using method and compositions of the present invention include but are not limited to melanoma, metastases, adenocarcinoma, thymoma, lymphoma, sarcoma, lung cancer, liver cancer, colon cancer, non-Hodgkins lymphoma, Hodgkins lymphoma, leukemias, uterine cancer, breast cancer, prostate cancer, ovarian cancer, cervical cancer, bladder cancer, kidney cancer, pancreatic cancer and the like.

CANCER

In one preferred embodiment the use of a target antigen from a tumour associated target antigen (TAA) in combination with the CTB of the present invention allows for the development of targeted antigen-specific vaccines for cancer therapy. The administration of composition comprising a TAA and optionally an immunomodulatory molecule, such as CTB, provides a powerful system to elicit a specifically enhanced cell mediated immune (CMI) response in terms of prevention in a host subject with an increased risk of cancer development (preventive immunisation), prevention of disease recurrence after primary surgery (anti-metastatic vaccination), or as a tool to expand the number of cytotoxic T lymphocyte (CTL) or CD8+ cells *in vivo*, thus improving their effectiveness in eradication of diffuse tumours (treatment of established disease). The CTB alone or a composition comprising the CTB as described herein may be used to elicit an enhanced CMI response in a host subject by the treatment of cells *ex vivo* prior to being transferred back to the tumour bearer (also known as adoptive immunotherapy).

The CTB as described herein or a composition comprising the CTB may be administered into a host subject either prior to any evidence of cancers such as melanoma (= preventative vaccination) or to mediate regression of the disease in a mammal afflicted with a cancer such as melanoma (therapeutic or immunotherapeutic vaccination).

DISEASES

The CTB as described herein or a composition comprising the CTB as described herein may be used to prevent and/or treat an enterotoxigenic disease. Examples of an enterotoxigenic disease includes but is not limited to cholera and enterotoxigenic

disease. By way of explanation, clinical cholera is an acute diarrhoeal disease that results from an oral infection with the bacterium *V. cholerae*. Cholera remains an important cause of illness in many developing countries and has been estimated to result in more than 200,000 deaths each year. In addition, diarrhoea caused by enterotoxigenic *E. coli* (ETEC) is an important cause of morbidity and mortality in many developing countries. Infection with enterotoxigenic *E. coli* (ETEC) is the most frequent cause of diarrhoea in the developing world and amongst travellers; it is responsible for more than one billion diarrhoeal episodes and one million deaths annually. Infection with ETEC is also an important cause of disease in animals. *V. cholerae* of serogroup 01 and enterotoxigenic *E. coli* may induce diarrhoea when multiplying in the gut of infected individuals by releasing cholera toxin (CT) or heat-labile enterotoxin (LT) respectively. Compositions comprising the CTB are advantageous in that the CTB component also provides for cross-protection against ETEC bacteria because the CTB molecule shares about 85% sequence homology with the *E. coli* LT toxin at the amino acid level and is the same length (103 amino acids). Without wishing to be bound by theory, there is evidence that both antibacterial and antitoxic immunity imparted by the B subunit of cholera toxin contributes to protection against both cholera and enterotoxigenic *E. coli*-caused diarrhoea and that an effective vaccine may evoke both anti-colonisation and anti-toxic immune responses in the intestine. Some anti-cholera and anti-ETEC vaccines have been developed which contain a combination of bacterial cell and toxin-derived antigens. (such as, for example, the Dukoral ® product).

AUTO-IMMUNITY

Autoimmune diseases considered for treatment using CTB include but are not limited to rheumatoid arthritis, multiple sclerosis, encephalomyelitis (or other neuron demyelinating diseases), diabetes, irritable bowel syndrome, Crohns disease, systemic lupus erythematosus (SLE) and female antipaternal immuno-contraception. WO 01/27144 teaches that CTB may be useful in the treatment of autoimmunitv and the induction of tolerance. By way of example, autoantigen specific peripheral T cell tolerance may be induced by CTB-autoantigen conjugates. Peripheral tolerance suppresses autoinflammatory disease by down regulating Thl. CTB-autoantigens can even reverse existing inflammatory cell proliferation in animals. Reversal of inflammatory cell proliferation (Thl tolerization) was shown in autoimmune models for diabetes in NOD mice (insulin-CTB conjugates; Bergerot *et al.* PNAS 1997 94: 461014.). Additionally it was shown for rheumatoid arthritis (HSP 60, Haque *et al.* 1996 EurJImmunol 26: 2650-2656; collagen type 11 conjugates) and experimental autoimmune encephalomyelitis (EAE) in Lewis rats (Sun *et al.*, Myelin basic protein MBP-peptide CTB conjugate, 1996 PNAS 93: 7196-7201). These conjugates were fed as single doses to animals before and after induction of disease. The CTB as described herein may also be used to induce tolerance to infection. For example, tolerization may be induced by administering the Leshmania major produced immunodominant antigen LACK, which generates a CD4+ T cell dominant response and drives a Th2 response to associated antigens. Tl12 facilitates lesion development and leads to unresolved infection. LACK+ transgenic mice are tolerant of LACK and resolve infection. LACK fed animals are tolerized and resolve infection. (McSorley *et al*, Eur. J. Immunol. 1998,28: 42432.) showed that CTB-LACK, fed or given nasally, down regulated the T cell proliferative response to

LACK and allowed animals to resolve (fight off) disease more effectively. Additionally, responses to diseases associated with the development of acute autoimmunity such as Reiter's syndrome or Lyme disease can be treated with tolerizing conjugates.

CTB AS A TARGET ANTIGEN

The CTB subunit protein as described herein may be used as a target for producing antibodies, such as monoclonal or polyclonal, specific to the CTB subunit protein or a fragement thereof. It is well known in the art that the CTB sequence comprises one or more antigenic determinants/epitopes which are immunologically reactive and which are capable of stimulating the production of antibodies for use under in vitro or in vivo conditions.

As used herein, the term "antigenic determinant" as used herein refers to a site on an antigen which is recognised by an antibody or T-cell receptor. Typically, it is a short peptide derived from or as part of a protein antigen.

As used herein, "immunologically reactive" is defined as the capability of the natural, recombinant or synthetic target sequence of the present invention to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

As used herein, the term "CTB target sequence" refers to a CTB sequence or variant, homologoue, derivative or fragment thereof which comprises an antigenic determinant/epitope which is immunologically reactive and which provides a means for identifying the CTB target sequence in a test sample.

The methods for producing antibodies against the target CTB sequence are known in the art. For the purposes of this invention, the term "antibody", unless specified to the contrary, includes but is not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library. Such fragments include fragments of whole antibodies which retain their binding activity for a target substance, Fv, F(ab') and F(ab')2 fragments, as well as single chain antibodies (scFv), fusion proteins and other synthetic proteins which comprise the antigen-binding site of the antibody. Furthermore, the antibodies and fragments thereof may be humanised antibodies. Neutralizing antibodies, i.e., those which inhibit biological activity of the substance polypeptides, are especially preferred for diagnostics and therapeutics. Antibodies may be produced by standard techniques, such as by immunisation with the substance of the invention or by using a phage display library. If polyclonal antibodies are desired, a selected mammal (e.g., mouse, rabbit, goat, horse, etc.) is immunised with an immunogenic polypeptide bearing a epitope(s) obtainable from an identified agent and/or substance of the present invention. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminium hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (Bacilli Calmette-Guerin) and Corynebacterium parvum are potentially useful human adjuvants which may be employed if purified the

substance polypeptide is administered to immunologically compromised individuals for the purpose of stimulating systemic defence.

Serum from the immunised animal is collected and treated according to known procedures. If serum containing polyclonal antibodies to an epitope obtainable from an identifed agent and/or substance of the present invention contains antibodies to other antigens, the polyclonal antibodies can be purified by immunoaffinity chromatography. Techniques for producing and processing polyclonal antisera are known in the art. In order that such antibodies may be made, the invention also provides polypeptides of the invention or fragments thereof haptenised to another polypeptide for use as immunogens in animals or humans.

Monoclonal antibodies directed against epitopes obtainable from an identifed agent and/or substance of the present invention can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal antibody-producing cell lines can be created by cell fusion, and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. Panels of monoclonal antibodies produced against orbit epitopes can be screened for various properties; i.e., for isotype and epitope affinity.

Monoclonal antibodies to the substance and/or identified agent of the present invention may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique originally described by Koehler and Milstein (1975 Nature 256:495-497), the human B-cell hybridoma technique (Kosbor et al (1983) Immunol Today 4:72; Cote et al (1983) Proc Natl Acad Sci 80:2026-2030) and the EBV-hybridoma technique (Cole et al (1985) Monoclonal Antibodies and Cancer Therapy, Alan R Liss Inc, pp 77-96). In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison et al (1984) Proc Natl Acad Sci 81:6851-6855; Neuberger et al (1984) Nature 312:604-608; Takeda et al (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies (US Patent No. 4,946,779) can be adapted to produce the substance specific single chain antibodies.

Antibodies, both monoclonal and polyclonal, which are directed against epitopes obtainable from an identifed agent and/or substance of the present invention are particularly useful in diagnosis, and those which are neutralising are useful in passive immunotherapy. Monoclonal antibodies, in particular, may be used to raise anti-idiotype antibodies. Anti-idiotype antibodies are immunoglobulins which carry an "internal image" of the substance and/or agent against which protection is desired. Techniques for raising anti-idiotype antibodies are known in the art. These anti-idiotype antibodies may also be useful in therapy. Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi et al (1989, Proc Natl Acad Sci 86:

3833-3837), and Winter G and Milstein C (1991; Nature 349:293-299).

Antibody fragments which contain specific binding sites for the substance may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse WD et al (1989) Science 256:1275-128 1).

DIAGNOSTIC KITS

In one embodiment, the present invention also includes a diagnostic composition or diagnostic methods or kits for (i) the detection and/or measurement of a target sequence, such as a CTB sequence in biological fluids and/or culture media. The kit of the present invention may also comprise reference means. As used herein, the term "reference means" refers to reference samples and/or control samples. The diagnostic compositions and/or methods and/or kits may be used in the following techniques which include but are not limited to; competitive and non-competitive assays, radioimmunoassay, bioluminescence and chemiluminescence assays, fluorometric assays, sandwich assays, immunoradiometric assays, dot blots, enzyme linked assays including ELISA, microtiter plates, antibody coated strips or dipsticks for rapid monitoring of urine or blood, immunohistochemistry and immunocytochemistry.

Assay methods for the detection of CTB subunit protein of the present invention may be tailored to detect the presence of the CTB subunit *in vitro* or to evaluate the efficacy of a particular therapeutic treatment regime or may be used in animal studies, in clinical trials, or in monitoring the treatment of an individual patient. In order to provide a basis for the progression or alleviation of a disease or disorder, a normal or standard profile for target expression should be established. Finally, the assay may be repeated on a regular basis to evaluate whether the values progress toward or return to the normal or standard pattern. Successive treatment profiles may be used to show the efficacy of treatment over a period of several days or several months.

SCREENS

Any one or more of the target sequences of the present invention may be used for identifying an agent capable of modulating the CTB target sequence or variants, homologoues, derivatives or fragments thereof in any of a variety of drug screening techniques. The target employed in such a test may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The abolition of target activity or the formation of binding complexes between the target and the agent being tested may be measured. Techniques for drug screening may be based on the method described in Geysen, European Patent Application 84/03564, published on September 13, 1984. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with a suitable target or fragment thereof and washed. Bound entities are then detected - such as by appropriately adapting methods well known in the art. A purified target can also be coated directly onto plates for use in a drug screening techniques. Alternatively, non-neutralising antibodies can be used

to capture the peptide and immobilise it on a solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralising antibodies capable of binding a target specifically compete with a test compound for binding to a target. Another technique for screening provides for high throughput screening (HTS) of agents having suitable binding affinity to the substances and is based upon the method described in detail in WO 84/03564. [It is expected that the assay methods of the present invention will be suitable for both small and large-scale screening of test compounds as well as in quantitative assays. Thus, the present invention relates to a method of identifying agents that modulate the target CTB sequence or variants, homologoues, derivatives or fragments thereof, the method comprising contacting a suitable target with the agent and then measuring the activity and/or levels of expression of the target sequence.

The present invention also relates to a method of identifying agents that selectively modulate the target sequence the method comprising contacting a suitable target with the agent and then measuring the activity and/or levels of expression of the target sequence. Test agents capable of modulating the activity of the CTB target sequence or variants, homologoues, derivatives or fragments thereof may be screened in assays which are well known in the art. Screening may be carried out, for example in vitro, in cell culture, and/or in vivo. Biological screening assays may be based on but not limited to activity-based response models, binding assays and bacterial, yeast and animal cell lines (which measure the biological effect of an agent in a smooth muscle cell, such or a tissue extract comprising same). The assays can be automated for high capacity-high throughput screening (HTS) in which large numbers of compounds can be tested to identify compounds with the desired target sequence modulating activity (see, for example WO 84/03564). Once an agent capable of modulating the target sequence activity - has been identified, further steps may be carried out either to select and/or to modify compounds and/or to modify existing compounds, to improve the target sequence modulation capability.

BACTERIAL VACCINE

The bacterial expression system itself may also be used as a vaccine for the prophylactic and therapeutic treatment of cholera and optionally other infectious diseases, especially in cases where the used strain has been engineered to express foreign proteins. By way of explanation, it has been shown that invasive bacteria (bacteria which can enter host cells), may be used to deliver vaccines. Preferably, in such embodiments the carrier bacterium in the vaccine is the invasive bacterium and expression of a heterologous nucleotide sequence is driven by a eukaryotic promoter. In such embodiments of the invention, expression of the heterologous nucleotide sequence will typically occur inside the cells of the vaccinated individual. Invasive bacteria include but are not limited to bacteria that are naturally capable of entering the cytoplasm or nucleus of animal cells, as well as bacteria which are not naturally invasive but which have been genetically engineered to enter the cytoplasm or nucleus of animal cells.

In one embodiment of the present invention, a *V. cholerae* host cell strain expression system of the present invention may be used as an immunising component in a

vaccine. In this regard, the attenuated *V. cholerae* strain lacking the functionality of the *thyA* gene and/or the expression plasmid comprising a functional *thyA* gene may be used as a carrier for a one or more heterologous antigens. Thus, the *V. cholerae* host cell strain comprising the expression system may also be used either alone or in combination with other components for prophylactic and therapeutic treatment of enterotoxin-induced illness in an individual (animal or human), such as cholera, ETEC and other enterotoxigenic-induced illnesses.

In one embodiment, the *V. cholerae* host cell expresses at least one heterologous antigen.

In another embodiment, the *V. cholerae* host cell expresses a number of different antigens so that the *Vibrio choleae* host cell is multivalent.

Accordingly, further aspects of the invention may include:

A *V. cholerae* host cell for use in the prevention and/or treatment of enterotoxin-induced illness, preferably a cholera and/or an ETEC induced illnesses.

The use of a *V. cholerae* host cell in the manufacture of a medicament in the treatment of an enterotoxin-induced illness, preferably a cholera and/or an ETEC induced illnesses.

A method of vaccinating a host mammalian subject against an enterotoxin-induced illness, preferably a cholera and/or an ETEC induced illnesses wherein the method comprises administering to the mammal a *V. cholerae* expression system as described herein or a vaccine comprising the *V. cholerae* expression system as described herein.

EXAMPLES

The present invention is also described by means of examples, including the particular Examples presented here below in which reference is made to the following Figures. The use of such examples anywhere in the specification is illustrative only and in no way limits the scope and meaning of the invention or of any exemplified term.

Brief Description of the Figures

Figure 1 shows the cloning of a 1.4 kb *Eco*RI/*Hind*III fragment in pUC19;

Figure 2 shows the insertion of a KanR-resistance gene block in the PstI site of the *V. cholerae thyA* gene in pUC19;

Figure 3 shows the PCR primers used to generate *thyA*-Kan fragment with XbaI ends;

Figure 4 shows the insertion of the *thyA* Kan fragment into XbaI restricted pNQ 705;

Figur 5 shows the elimination of the start of the coding region of the Kanamycin gene and part of the *thyA* gene;

Figure 6 shows the insertion of the Δ*thyA* ΔKan fragment into XbaI restricted pDM4;

Figure 7 shows the PCR amplification and subcloning of the *E. coli thyA* gene in pUC19;

Figure 8 shows the generation of pMT-*thyA*/cat;

Figure 9 shows a comparison of the signal-peptide mature protein region in *ctxB* from pMTctxB(*thyA*)-2, native *ctxB* and native *eltB*;

Figure 10 shows the insertion of the *eltb-ctxB* coding fragment from pML-LCTBλ2 in pMT-*thyA*/cat;

Figure11 shows the insertion of tac promotor in pMT-*thyA*/cat(ctxB) and the generation of pMT-ctxB/*thyA*(cat);

Figure 12 shows the removal of the cat gene, generation of pMT-ctxB/*thyA*;

Figure 13 shows the PCR reaction to remove superfluous *V. cholerae* DNA from pMT-ctxB/*thyA*, generation of pMT-ctxB*thyA*-2;

Figure 14 is a graphic representation of the parts of pMT-ctxB*thyA*-2 that has been sequenced on Master Seed lot and consistency batches;

Figure 15 shows an amino acid comparison of native CTB, rCTB 401 and LTB from *E. coli*;

Figure 16 presents the DNA sequence of the expression plasmid pMT-ctxB*thyA*-2;

(204-295: E. coli thyA coding region; 1192-1876: Col E1 origin of replication; 2339-2710: eltB-ctxB coding region; 2402-2710: ctxB coding region; and 2732-2759: trpA terminator)

Figure 17 shows the insertion of the eltb fragment in pmmb66eh; generation of pmmb68;

Figure 18 shows the generation of the pJS162 plasmid;

Figure 19 shows the generation of the pJS752-3 expression vector;

Figure 20 shows the detailed figure of the pJS752-3 plasmid;

Figure 21 shows the PCR fragment generated for sequencing of the agctb gene;

Figure 22 show the suggested transcriptional stop for the *agctb* gene. The –G value is -19.1 kcal/mol;

Figure 23 shows the open reading frames of the agctb 1201 bp fragment;

(Potential open reading frames were read in all six different readning frames (1, 2, 3 and -1, -2 and -3). Complete bars indicate stop codons, short bars indicate potential start (ATG) codons. The agctb gene is encoded in the first long open reading frame in 1)

Figure 24 shows the alignment of the amino acid sequence of rCTB, CTB from classical biotype and from El Tor biotype;

Figure 25 Complete nucleotide sequence of pJS752-3;

(Underline = sequenced in master seed lots and consistency batches.
Bold and italic = Open reading frame of Eltb leader sequence and ctxB gene = agctb gene)

Figure 26 Alignment of the *agctb* gene and flanking regions in *V. cholerae* strain 213 from six different 1000 liter fermentations and two Master Seed Lot preparations;

Figure 27 is a DNA alignment of *ctxB* coding and flanking regions from pMT-ctxB*thyA*-2 in Master seed lots and consistency batches; and

Figure 28 shows a comparison between the sequences of LTB from a human ETEC isolate (Leong, et al., 1985 Infect Immun 48; 73-78) and CTB (M. Lebens, unpublished). For differences resulting in amino acid changes, the residue at the corresponding position of LTB is given above the DNA sequence. Numbers below the CTB amino acid sequence indicate the position of amino acid residues in the mature CTB and LTB proteins.

The one-letter amino acid symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission are provided below together with the three-letter codes which are also provided for reference purposes.

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Table 1: Some CTB/TLB sequences

Protein	NCBI Accession No
CTB	GI: 209555
	GI: 433859
	GI: 48420
	GI: 48888
	GI: 155296
	GI: 48347
	GI: 758351
	GI: 1827850
	GI: 808900
	GI: 229616
	GI: 998409
	GI: 2144685
	GI: 1421511
CTB classic (596B)	GI: 48890
CTB Ogawa 41	GI: 2781121
CTB Ogawa 41 (R35D)	GI: 1421525
Classic LTB	GI: 3062900
	GI: 1169505
	GI: 1395122
	GI: 145833
LT 87	GI:1648865
	GI: 223254
	GI: 408996
	GI: 494265
	GI: 69630
LT-IIa	GI: 146671
LT-Iib	GI: 152784

Table 2:Some Leader/Signal sequences

Endotoxin Signal Sequece	Signal sequence
LTB signal sequence	MNKVKCYVLFTALLSSLCAYG
CTB V. cholera classic strain 569B CTB gene signal sequence	MNKVKFYVLFTA LLSS LCAH GAPGYAHG
LTB Signal sequence for "213" strain	MNKVKFYVLFTA LLSS LCAH GAPGYAHG = 28aa
LTB Signal sequence for "401" strain	MNKVKFYVLFTA LLSS LCAH G = 21aa

5 **Table 3: Comparison between known CTB plasmids and the CTB plasmid described herein**

Plasmid	Reference	Host cell	Plasmid Selection Marker	Plasmid size	CTB sequence	Yield of CTB
PJS162 (pJS213)	Sanchez and Holmgren (1989)	Toxin deleted JBK70 *V. cholerae*	Ampicillin resistance marker (AmpR)	About 10.2kb	LTB leader sequence CTB coding sequence CTB genomic sequence (down stream of CTB gene)	0.04-0.05mg/ml or 0.05-0.1mg/ml (see p482, col 2)
PML358	Lebens *et al* (1993)	Rifampicin Resistant CtxA deleted derivative of classical 569B *V. cholerae* JS1569 strain	Ampicillin resistance marker (AmpR)	Plasmid size not disclosed	LTB leader sequence CTB coding sequence CTB genomic sequence (down stream of CTB gene)	1mg/ml when ampicillin was maintained in the growth medium
PNU212-CTB	Ichikawa *et al* (1993) FEMS Microbiol Lett 111: 219-224	Bacillus Brevis	Erythromycin resistance gene (EmR)	About 4.8kb	Promoter and signal sequence for one of the major extracellular proteins (MWP) of B Brevis used with CTB sequence	1.4mg/ml after a period of 4-5 days with antibiotics present in the growth medium
PML-CTBtac1	WO 01/27144 (Active Biotech AB) (page 46-47 of app as filed and Figure 2)	Bacterial host strain	Kanamycin resistance marker	About 3.66kb	CTB leader sequence CTB coding sequence CTB genomic sequence (down stream of CTB gene)	Five times the product generated by pJS162 (see page 46, l32-34) (0.8mg/ml – see page 47, lines 12-13)
PJS752-3	Figures 19 and 20 as described herein	*V. cholerae*	Ampicillin resistance marker (AmpR)	About 5.75kb	LTB leader sequence CTB coding sequence CTB genomic sequence (down stream of CTB gene)	About 0.4mg/ml (see yield table)
PMT-ctxB*thyA*-2	Figures 13 and 14 as described herein	*V. cholerae*	*thyA* (non antibiotic selection marker)	About 2.8kb	LTB leader sequence CTB coding sequence (about 800bp removed downstream of CTB coding sequence)	About 1.4mg/ml (see yield table) after only 18 hours

Example I

A. Source materials

The origin of the *ctxB* gene is the *V. cholerae* serotype O1 strain 395 (Ogawa) [2]. The *eltB* signal sequence was obtained from plasmid pMMB68 [3]. The ligation of the *eltB* and *ctxB* gene is described in [4].

The origin of replication is ColE1 obtained from pBlueScript KS$^-$ (Stratagene).

The origin of the *tac* promotor used in pMT-ctxB*thyA*-2 is from plasmid pKK223-3 (Pharmacia).

The DNA sequence used to PCR amplify the *E. coli thyA* gene was *E. coli* SY327 [5].

The Kan^R resistance gene block used in inactivation of the chromosomal *V. cholerae thyA* locus was obtained from pUC4K (Pharmacia).

The suicide vectors used for site-directed mutagenesis of the *V. cholerae thyA* locus were pNQ705 [6] and pDM4 described in [7].

The sequence of the *V. cholerae thyA* gene was determined at SBL Vaccin AB and is published in EMBL/Genebank under accession No AJ006514.

A.1 Construction of the *V. cholerae* Inaba strain 401 Classical biotype for production of rCTB.

A.1.2. Construction of the host strain V. cholerae JS1569 ΔthyAΔKan.

The *V. cholerae* strain JS1569 Δ*thyA*Δkan is a classical O1 cholera strain originally derived from *V. cholerae* strain 569B; ATCC No 25870. The two copies of the cholera enterotoxin genes have been replaced in this strain by attenuated *ctx* genes from *V.cholerae* strain 395 that have been modified by site-directed mutagenesis [2, 8]. The attenuation comprises of a deletion of the cholera toxin A subunit gene.

The *V.cholerae* strain 395 *ctx* genes were cloned into the plasmid pCVD15 as a *Hind*III fragment. A 550 base-pair *Xba*I-*Cla*I restriction fragment was excised, deleting the A1 gene at the position encoding the 10:th amino acid and further into the gene A2 encoding the first amino acid of the A2 peptide. The resulting plasmid pCVD30 was digested with *Hind*III and the *Hind*III fragment cloned in the mobilisation vector pJBK85 for mobilisation into the *V.cholerae* strain 569B yielding strain CVD 103 [8]. In all 94% of the *ctx*A gene was deleted [2, 8 and SBL Vaccin test method PT00025 which method is a Polymerase Chain Reaction (PCR) with primers designed so that the deletion in the ctxA gene can be detected in both ctxA loci simultaneously (a multiplex PCR)]. To render this strain suitable for conjugation experiments a rifampicin resistant derivative of CVD103 has been selected, the JS1569 strain [9]. The deletion and insertional inactivation of the *thyA* gene is described below.

A.1.3. Inactivation of the thyA gene in V. cholerae JS1569.

In the course of isolation and sequencing of the *thyA* gene from JS1569 [Carlin *et al*, Genebank accession no AJ006514) an *Eco*RI-*Hind*III fragment encompassing the entire *thyA* gene was cloned in the vector pUC19 on a 1.4 kb DNA fragment. This plasmid is called *thyA* 1.4 (Fig 1.).

A.1.4 Inactivation of the thyA gene by insertion of a Kan^R gene block.

The plasmid p*thyA*1.4 was cleaved with *Pst*I, and ligated to a *Pst*I fragment of the Kan^R gene block from pUC4K (Pharmacia Biotech). The ligation mixture was electroporated into *E. coli* HB101 and transformants were selected for by plating on

Syncase agar plates containing ampicillin and kanamycin. This plasmid is called p*thyA* **Kan**.(Fig. 2).

The *thyA* Kan^R gene was PCR amplified from p*thyA* Kan using a mix of *Taq* and *Pwo* DNA polymerases which yields PCR fragments with high sequence fidelity (Expand™ High fidelity PCR system, Boehringer Mannheim) The primers used were *thyA*-10 GCT **CTA GAG** CCT TAG AAG GCG TGG TTC and *thyA*-11 GCT **CTA GAG** CTA CGG TCT TGA TTT ACG GTA T generating a PCR fragment with *Xba*I ends (Fig 3).

This fragment was digested with *Xba*I and ligated into the vector pMAL-C2 that had been digested with *Xba*I and dephosphorylated as indicated above. The fragment size and orientation was confirmed by restriction enzyme analysis.

A.1.5 Insertional inactivation of the thyA gene in the V. cholerae chromosome by site-directed mutagenesis.

The suicide vector pNQ705 [6] (Fig 4) contains the R6K origin of replication and hence has to be maintained in a host harbouring the *pir* gene. It also contains the *mobRP4* genes and a CAT gene allowing for chloramphenicol selection.

The *thyA* Kan^R gene was excised from pMAL-C2 as a *Xba*I fragment and ligated into *Xba*I digested pNQ705 (Fig 4.) The ligation mix was electroporated into *E.coli* SY327 [*Δ(lac pro) argE*(Am) *rif malA recA56*] and transformants were selected for on plates containing chloramphenicol.Restreaked individual colonies were analysed with restriction enzymes for presence and orientation of insert.

The resulting plasmid pNQ705 *thyA* Kan^R was transformed by electroporation into *E. coli*S17-1 (*thi pro hsdR hsdM*$^+$ *recA* RP4-2-Tc::Mu-Km::Tn7).

The mating between JS1569 *thyA*$^-$ (a trimetoprim resistant variant of JS1569 carrying a single point mutation in the *thyA* gene [10] and *E. coli* S17-1 (pNQ705 *thyA* Kan^R) were done as streak matings on LB agar supplemented with rifampicin (50 μg/ml), Thymine (200 μg/ml) and Kanamycin (50μg/ml) at 37°C. Individual colonies arising from the conjugation were transferred to liquid LB broth with rifampicin, thymine and Kanamycin supplements as above and were passaged for three days in this medium.

Transconjugants were at this time tested in PCR for insertion of the *thyA* Kan^R gene. Cultures that had the expected PCR fragment were plated out on LB agar supplemented as above.

Individual colonies were now picked and tested for sensitivity to chloramphenicol (25μg/ml) and resistance to Kanamycin. These colonies were restreaked and individual colonies frozen.

Phenotypically this strain was thymine dependent for growth.

A.1.6. Insertional inactivation of the Kan^R gene and deletion of the thyA gene.
An experiment was designed to replace the functional Kan^R gene with a truncated nonfunctional version and to remove a substantial part of the *thyA* gene to further ensure the stability of the thymine dependence in this strain. For this experiment the *thyA* Kan fragment with *Xba*I ends was subcloned into *Xba*I digested pNEB193 (New England Biolabs). Two PCR primers were designed *thyA*-14 and *thyA*-15 that both included *Xho*I cleavable ends. The primers were designed so as to eliminate 209 basepairs in the *thyA* gene and to eliminate 144 basepairs in the Kan^R gene (totalling 266 bp from the Kan^R gene block. The deletion in the Kan^R gene included the first 48 aminoacids of the Kan^R gene which would make succesful transconjugants Kan^S. The resulting PCR fragment was cleaved with *Xho*I and allowed to selfligate, transformed into *E.coli* and transformants were selected for on ampicillin containing agar. Colonies were tested for Kanamycin sensitivity and the deletion was confirmed with restriction enzyme analysis.

From this plasmid the resulting Δ*thyA* ΔKan fragment was excised as an *Xba*I fragment. This fragment was inserted into the suicide vector pDM4 that had been digested with *Xba*I. pDM4 is derived from pNQ705 by replacing the multicloning site and insertion of the *SacB* gene from *Bacillus subtitlis*. The *SacB* gene encodes levansucrase gene that is lethal to Gram negative bacteria. The ligation mixture was transformed into *E. coli* SY327 and transformants were selected for by chloramphenicol. The insert size and orientation was verified by restriction enzyme analysis. For mating experiments the plasmid pDM4 Δ*thyA*ΔKan was transformed into *E. coli* S-17. Mating was performed between *E. coli S-17* (pDM4 Δ*thyA* Δ Kan) and JS1569 Δ*thyA* Kan on LB plates containing rifampicin, chloramphenicol and thymine. Transconjugants were grown further in LB broth with rifampicin, chloramphenicol and thymine. After passaging for three days in this medium colonies were plated on LB plates containing thymine and 5% sucrose. Emerging colonies on these plates were tested by replica plating for growth on medium with and without thymine, thymine plates containing chloramphenicol and plates containing kanamycin. Chloramphenicol and kanamycin sensitive colonies with a requirement for thymine were selected and tested in PCR with appropriate primers for a replacement of the *thyA* Kan^R gene block with the Δ*thyA* ΔKan insert. Single colonies of this strain were restreaked. For these colonies re- confirmation of the genotype (rifampicin resistance, deletion of *ctxA* loci, thymine dependence, chloramphenicol, and kanamycin sensitivity) was done and the strain was namned JS1569 Δ*thyA* ΔKan.

Further characterisation involved PCR amplification and partial sequencing of the modified chromosomal *thyA* locus in this strain. It was found that the point mutation in the trimetoprim resistant *thyA*$^-$ strain used for the first mating experiment had been changed to wildtype *i.e.* the thymine dependence of this strain is caused by the deletion of the *thyA* gene further downstream. DNA sequencing also confirmed the deletion of the *thyA* and Kan gene block (data not shown).

B. The expression plasmid pMT-ctxB*thyA*-2

B.1.1. Cloning of the E. coli thyA gene.

For the cloning of the *E. coli thyA* gene the published sequence (Genebank accession no J01709) was used to design PCR primers

ML*thyA*-1: $^{5'}$ GGG GG*C TCG AG***G TTT GTT CCT GAT TGG TTA CGG**$^{3'}$

Letters in bold indicate sequence from the published sequence (bases 16-39 on the sense strand), italic letters indicate a *Xho*I site added to the sequence.

ML*thyA*-2: $^{5'}$ GGG GG*G TCG AC***G TTT CTA TTT CTT CGG CGC ATC TTC**$^{3'}$

Letters in bold indicate sequence from the published sequence (bases 1152-1128 on the non-sense strand), italic letters indicate a *Sal*I site added to the sequence.

These primers were used to amplify the *thyA* gene from *E. coli* SY327.

The resulting PCR fragment was blunt-end repared with T4 polymerase and cloned into pBluescript KS$^-$ (Stratagene) in the *Eco*RV site in this vector. The ligated plasmid was transformed into *E. coli* XL1-Blue (Stratagene). Transformants were selected for on LB plates supplemented with ampicillin on the basis of blue/white colonies in the presence of X-Gal and IPTG. The inserted fragments size and orientation was confirmed by restriction enzyme cleavage. The functionality of the *thyA* gene was confirmed by electroporating the recombinant plasmid into JS1569 *thyA*$^-$ and selection both for ampicillin resistance and growth on modified syncase medium in the absence of thymine. This plasmid is called pML-*thyA*(XS).

Sequencing of the pMT ctxB(*thyA*-2) plasmid at a later stage (pre Master and Master seed lot as well as end of production cells) indicate that the *Sal*I site coded for on primer Ml*thyA*-2 has been lost (most likely at this stage of development).

Instead of the expected sequence (sense strand):

$^{5'}$ GAA ACG TCG A<u>CC CCC C</u>AT CAA GCT TAT$^{3'}$

the following sequence was found:

$^{5'}$ GAA ACG TCG AAT CAA GCT TAT$^{3'}$

i.e. the underlined sequence CC CCC C has been lost in this procedure, this sequence are the additional 6 G:s added to the 5' end of primer Ml*thyA*-2. The reasons for this are unclear, however the deletion is not in the coding region for any protein and has been found both in Master Seed lots and in end of production cells and is thus stabily maintained.

B. 1.2. Generation of a cloning vector carrying the E. coli thyA gene.

For the cloning of the *E. coli thyA* gene the vector pML-X1 was used (Fig 8.) This plasmids origin is pBC SK$^-$ (Stratagene). In pML-X1 the origin of replication (ColE1) is flanked by unique *Bgl*II and *Stu*I sites and it also carries the chloramphenicol (cat) gene from pBC SK$^-$ (Fig 8).

pML-X1 DNA was digested with *AgeI*/*Stu*I restriction enzymes and blunt-end repaired with T4 polymerase.

The pML-*thyA*(XS) vector was digested with *Bam*HI/*Sal*I and also blunt-end repaired (Fig 8.) The two DNA preparations were mixed and ligated. The ligation mix was electroporated into JS1569 4.4 (A trimetoprim resistant variant of JS1569 carrying a single point mutation in its *thyA* gene) and thymine dependence and chloramphenicol reistance were selected for. The resulting plasmid pMT-*thyA*/cat was restreaked and subjected to extensive restriction enzyme analysis to verify the size and orientation of its different components (Fig 8.)

B.1.3. Insertion of* ctxB *into pMT-thyA/cat.

In order to insert the desired *ctxB* gene into the pMT-*thyA*/cat plasmid a 1.2 kb *Eco*RI/*Xho*I fragment from plasmid pML-LCTBλ2 was obtained. The plasmid pML-LCTBλ2 has been described earlier [4], briefly the *ctxB* gene was isolated from the plasmid pCVD30 [2]. The *ctx*B gene contained in the pCVD30 plasmid [2] originates from the *V. cholerae* serotype O1 strain 395 (Ogawa).

The *ctxB* gene in pML-LCTBλ2 is upstreams fused to the *eltB* signal peptide from the heat-labile enterotoxin of *E. coli* in such a way that a naturally occurring *Sac*I site could be used (Fig 9.). This also introduces an Alanine as the N-terminal amino acid rather than the naturally occurring Threonine. This modification of the N-terminal sequence and signal peptide has led to that only a single N-terminal sequence is formed from this new expression plasmid for rCTB as compared to the previously used pJS752-3 (see section C.2 below). Downstream of the *ctxB* gene on the *Eco*RI/*Xho*I fragment from pML-LCTBλ2 the powerful *trpA* terminators are located, effectively terminating m-RNA transcription. The *Eco*RI/*Xho*I fragment from pML-LCTBλ2 was ligated into the pMT-*thyA*/cat plasmid that had been digested with the same enzymes (Fig 10), resulting in the plasmid pMT-*thyA*/cat(ctxB), which lacks a promotor upstream of the *eltb* signal peptide.

B.1.4. Insertion of the* tac *promotor into pMT-thyA/cat(ctxB).

The *tac* promotor was inserted as a 256 base-pair *Bam*HI/*Eco*RI fragment originally obtained from the cloning vector pKK223-3 (Pharmacia) (Fig 11). Ligated DNA from this reaction was introduced into *V. cholerae* JS1569 4.4 and colonies were selected for on the basis of growth in the absence of thymine and resistance to chloramphenicol. Transformants were screened both by restriction enzyme analysis of the recombinant plasmids and the production of CTB. Single colonies with the highest rCTB production as judged by GM1-ELISA were selected. The recombinant plasmid was named pMT-ctxB/*thyA*(cat).

B.1.5. Removal of the cat gene.

To remove the cat gene *i.e.* to obtain an expression plasmid without any antibiotic selection marker the pMT-ctxB/*thyA*(cat) plasmid was digested with the restriction enzymes *Bam*HI and *Bgl*II. The cut plasmid was religated and again electroporated into the *V. cholerae* strain JS1569 4.4. Transformants were selected on the basis of growth in the absence of thymine and in the absence of chloramphenicol. Indivudal colonies were screened for sensitivity to chloramphenicol and checked for presence of plasmid in Wizard Miniprepps. Plasmids were analysed with restriction enzymes.

Culture supernatants from these colonies were subjected to GM1 ELISA. The resulting plasmid was pMT-ctxB/*thyA* (Fig 12).

B.1.6. Removal of superfluous V. cholerae DNA from pMT-ctxB/thyA, generation of pMT-ctxBthyA-2.

The *Eco*RI/*Xho*I fragment from pML-LCTBλ2 consists of approx 1200 base-pairs, of these only about 400 base-pairs code for the *ctxB* gene. There is in one readingframe going in the other direction of *ctxB* an open reading frame possibly coding for an *orfF* protein in the *pyrF* operon. The sequence is incomplete and thus probably not expressed. In order to remove the non-coding CTB portion, PCR primers were designed, the first so as to include the end of the *ctxB* gene (in italics below) and a *Spe*I site (in bold below). The other PCR primer was designed to include the *trpA* terminators (in italics below) and also a *Spe*I site (in bold below).

CTB3' :$^{5'}$GGG GGA **CTA GT***T TAA TTT GCC ATA CTA ATT GCG GCA ATC G*$^{3'}$

TrpA term:$^{5'}$GGG GGA **CTA GTC** AAT TGA AGC TTA *AGC CCG CCT AAT GAG CG*^{3}

The pMT-ctxB/*thyA* plasmid served as template for the PCR reaction. After obtaining an PCR fragment of the correct size, this was gel-purified, and digested with *Spe*I. The plasmid was allowed to self-ligate and was electroporated into *V. cholerae* JS1569 4.4. Transformants were selected on the basis of of growth in the absence of thymine, single colonies isolated, restreaked, and plasmid DNA from these cultures was analysed by restriction enzyme analysis. Approximately 800 base-pairs of DNA including the entire *orfF* coding sequence was removed. GM1 ELISA indicated that this had no effect on the expression of CTB.

The resulting plasmid was called pMT-ctxB*thyA*-2 and is the final construct used in conjunction with the host strain *V. cholerae* JS1569 Δ*thyA*Δkan to form the rCTB producing strain *V. cholerae* strain 401.

B.1.7. Insertion of pMT-ctxBthyA-2 in V. cholerae JS1569 ΔthyAΔkan.

The plasmid preparation from pMT-ctxB*thyA*-2 in *V. cholerae* JS1569 4.4 was electroporated into *V. cholerae* JS1569 Δ*thyA*Δkan thereby forming the strain *V. cholerae* Inaba strain 401 classical biotype. Transformants were selected for by their ability to grow in the absence of thymine. Individual colonies were tested for their ability to produce rCTB by colony lifts on nitrocellulose filters (SBL test method PT00020 as described below) using a monoclonal antibody specific for both CTB and LTB (*E. coli* heat-labile enterotoxin). Colonies were re-streaked to obtain single colonies and cultures from these colonies were used for plasmid analysis and extensive restriction analysis and finally frozen.

SBL Test Method PT00020

This is a method which is used to distingusih colonies of the rCTB producing strain that can produce rCTB from those that have lost that capacity. The methodology used consists of growing the bacterial colonies to be tested on an agar plate, transferring the

colonies to a nitrocellulose filter. This filter is incubated with a monoclonal antibody specific for pentameric rCTB, washed and then incubated with an anti mouse IgG alkaline phosphatase conjugate. After washing the filters are dveloped with a precipitating dye, leaving the rCTB colonies bluish-black while non-producing colonies are left essentially colourless

C. Description of the *V. cholerae JS1569* Δ*thyA* Δkan strain carrying the pMT-ctxB*thyA*-2 expression vector: strain *V. cholerae* 401.

C.1. Nucleotide sequence of the ctxB gene and amino acid sequence of the translated polypeptide.

C1.1. Detailed nucleotide sequence of the ctxB gene and flanking regions.

Plasmid DNA was purified from CsCl gradient ultracentrifugation and sequenced The complete nucleotide sequence of plasmid pMT-ctxB*thyA*-2 is given in Figure 16. 95% of the plasmid has been sequenced (to be completed) at the stage before production of seed lots. In the Master Working Seed lot the coding region for the *ctxB* gene has been sequenced from two different tubes in the Master Seed lot Bank (as indicated in Fig. 14) (see Figure 27). End of production cells from three consistency lots has also been sequenced, Figure 27. The sequence obtained from both Master Seed lot as well as from consistency lots show 100% identity.

C.2. Amino- terminal sequence of the mature recombinant protein.

In Fig. 15 the amino acid sequence from the rCTB produced by strain *V. cholerae* 401 is compared with the amino acid sequences of native CTB and native LTB toxin from *E. coli*. As can be seen in Fig 15, the amino acid sequence of the signal peptide of LTB and rCTB 401 are identical as is the N-terminal amino acid of the mature protein. The comparison of native CTB (classical biotype) and rCTB 401 shows that the only difference is the N-terminal amino acid (Threonine in native CTB and Alanine in rCTB 401. This modification is justified by the previous experience with the rCTB 213 molecule, which also has the *eltB* signal sequence linked to the *ctxB* sequence. There were also four additional amino acids included in the sequence linkage region of this rCTB 213 molecule due to the methods for recombinant DNA technology available at the time [9].

Experience with the rCTB 213 in fermentor scale production revealed that up to six rCTB species with different amino termini could be isolated.
With this knowledge in mind the rCTB 401 linkage was designed so that the extra amino acids in the linkage region were removed plus that the N-terminal amino acid was replaced so as to be identical to that of native LTB *i.e.* an Alanine.
This modification has proven itself to be advantageous. There is only one N-termini in rCTB 401 isolated in all experimental and consistency batches of rCTB 401 irrespective of fermentation time and conditions.

C.3. Mode of expression.

The pMT-ctxB*thyA*-2 plasmid that harbours the *ctxB* gene does not contain the gene for the strong repressor (*lacI*^{q}). In *V.cholerae* it is not known if there is a repressor (*lacI*) in the genome. *V.cholerae* do not ferment lactose but have a *lacZ* gene [12]. It is reasonable to assume that the eventual repressor will not be as potent as *lacI*^{q} and

also that it would be present in much smaller quantities than the promotor (*tac*P) which is located on a high copy number plasmid. As a result the expression of *agctb* is in practice constitutive.

D. Stability of the expression system.

D.1. Storage stability.

The Master and Working Seed lots of the *V.cholerae* strain 401 have been stored for less than a year respectively at -65°C or colder. Genetic stability for both the Master and Working Seed lots have been demonstrated after 6 months of storage since 100% of the colonies produced rCTB when grown for production of consistency lots. Previous experience with the Seed lot system of the rCTB producing strain *V. cholerae* 213 indicate that for that strain stability is excellent for more than 7 years. The Mster and Working Seed lots are included in a stability testing program with testing every 5 years.

D.2. Stability in extended culture time.

In experiments designed to investigate the stability of the plasmid retention and rCTB production, the *V.cholerae* strain 401 has been grown on a shaker in Modified Syncase broth at 37°C. Every day the culture was diluted down 10.000 fold in fresh medium. This was commenced for 11 days. On day 7 and 11 the culture was spread on Modified Syncase agar and the ability of the colonies to produce rCTB was tested with a colony blotting technique [SBL Vaccin test method PT00020 as described below] using a monoclonal antibody specific for LTB and crossreacting with CTB. After more than 100 generations (11 days of growth) 100% of the colonies retained their capacity to produce rCTB.

SBL Test Method PT00020

As indicated above (see for example section B.1.7), PT00020 is a method which is used to distingusih colonies of the rCTB producing strain that can produce rCTB from those that have lost that capacity. The methodology used consists of growing the bacterial colonies to be tested on an agar plate, transferring the colonies to a nitrocellulose filter. This filter is incubated with a monoclonal antibody specific for pentameric rCTB, washed and then incubated with an anti mouse IgG alkaline phosphatase conjugate. After washing the filters are dveloped with a precipitating dye, leaving the rCTB colonies bluish-black while non-producing colonies are left essentially colourless

D.3. Production stability.

The production scale for the *V. cholerae* strain 401 is 500 litres. The medium is the same modified syncase medium as used above with the exception that glucose is used instead of sucrose. To investigate the plasmid retention, and to show consistency, samples were taken at break-point from three consecutive 500 litre production fermentations. After approx. 18 hours of growth at 37°C in the main 500 litre fermentor 100% of the cells have retained their capability of producing rCTB.

D.4. Stability of the genetic construct during production fermentation.
To show the stability of the genetic construct, DNA was prepared from breakpoint harvests of *V.cholerae* strain 401. The plasmid DNA was purified by CsCl ultracentrifugation and sequenced as outlined in Fig 14. (see also Figure 27). The first base in the consensus sequence corresponds to base No 2210 in pMT-ctxB*thyA*-2 DNA, the last base corresponds to base 220 in pMT-ctxB*thyA*-2. The sequenced region encompasses sequence before the *tac* promotor, the entire *eltB-ctxB* and ends 18 bases inside the coding region for the *thyA* gene. The sequence determined from samples taken during production fermentation show the identical DNA sequence of the *tac* promotor, the *eltb-ctxB* gene and flanking DNA as that obtained from seed lot.(Figure 27) thereby demonstrating the stability of the construct.

Background to the production of the Vibrio cholerae "213" strain
The cholera toxin (CT) is almost exclusively found in *V. cholerae* of serotype O1. Recently however, a new serotype of highly virulent *V.cholerae* has emerged; the serotype O139 which also produces CT. The cholerae holotoxin consists of two subunits; the A subunit (CTA) responsible for the activation of adenylate cyclase in the intestinal cells of the host and a pentamer of identical B subunits (CTB) that binds the holotoxin to its GM1 ganglioside receptor.
Genetically the cholerae toxin genes (*ctx*AB) are located in a "virulence cassette" together with the *cep*, *ace* and *zot* genes associated with colonisation and pathogenesis of *V.cholerae*. In *V.cholerae* O1 of classical biotype there are two copies of this virulence cassette at different locations on the chromosome, in strains of the El Tor biotype there is only one copy. In *V.cholerae* O139 there are at least two copies of the virulence cassette [1]. The *ctx*AB genes are read into one polycistronic messenger-RNA. Analysis of the DNA sequence indicate that there are presumably individual ribosome binding sites for both genes *i.e.* there is no translational coupling.
The recombinant B subunit (rCTB) of the cholera toxin (CT) is the only component of the oral cholera vaccine Dukoral® that is made with recombinant DNA techniques. The gene for rCTB is inserted in an expression vector in a *V.cholerae* O1 strain. The expression of the rCTB is designed so that rCTB is overproduced and accumulates in the growth medium. The strain harbouring the plasmid is a *V. cholerae* O1 of the Inaba serotype, classical biotype and is resistant to 50 µg/ml of ampicillin because of the presence of the pJS752-3 plasmid. The strain is deleted in both its chromosomal *ctx*A loci thereby abrogating the production of intact holotoxin. The designation of the production strain with the pJS752-3 expression vector is *V.cholerae* strain 213.

A. Source materials.
In order to establish a recombinant *V. cholerae* expression system for the production of cholera toxin B subunit (rCTB) a hybrid gene has been constructed: the "*agctb*" gene [14]. The *agctb* gene consists of two parts: the first part of the *elt*B gene encoding the leader peptide of the *E.coli* LTB protein fused with the second part; the *ctx*B gene of *V. cholerae*. The *ctx*B gene product is the B subunit of the cholera toxin which originates from the *V. cholerae* serotype O1 strain 395 (Ogawa) [15]. The agctb gene was introduced into the commercially available expression vector pKK223-3 (Pharmacia Biotech) to yield the expressionvector pJS752-3. In pJS752-3 the *agctb* gene is expressed under the control of the strong *tac*P promotor [14]. The pJS752-3 plasmid also harbours the ampicillin resistance gene as a selection marker.

A.1 Preparation of the V. cholerae strain 213 for production of rCTB

A.1.1. Construction of the agctb *gene*
The *agctb* gene consists of two parts; the leader sequence from the *E.coli* LTB gene (*elt*B) fused via a linker to the *ctx*B gene of *V.cholerae*.

A.1.2 Isolation strategy of the elt*B leader sequence and the* ctx*B gene.*
The DNA coding for the *elt*B leader sequence has been ligated as an *Eco*RI-*Hind*III fragment into plasmid pMMB66EH [16] yielding the plasmid pMMB68. This plasmid contains a *tac* promotor directly upstream of the *Eco*RI site (Figure 17) [17]. The *ctx*B gene contained in the pCVD30 plasmid [18] originates from the *V. cholerae* serotype O1 strain 395 (Ogawa). The generation of the hybrid *agctb* gene profits from the presence of the unique *SacI* site in position 78 of the DNA sequence encoding the LTB leader peptide, cleaving the leader sequence at the penultimate base (Figure 18a). The *ctx*B gene was excised from the pCVD30 plasmid [15] by cleavage with *NdeI* and *HindIII*. The recipient plasmid pMMB68 was cut with *SacI* and *HindIII* thereby removing the DNA sequence encoding the mature LTB protein (Fig 18b). In order to allow fusion of the *SacI* site of the *elt*B-derived sequence to the *NdeI* site of *ctx*B coding region, a synthetic linker was used. The linker had cohesive ends of *Sac*I and *Nde*I sites in its 5′ and 3′ end, respectively.

$$^{5'}CCC\ GGG^{3'}$$

$$^{3'}T\ CGA\ GGG\ CCC\ AT^{5'}$$

After ligation with T4 ligase, the vector pJS162 was generated. The following DNA sequence was generated at the junction between *eltB* and *ctxB*:

$^{5'}$GCA CAC GGA GCT *CCC GGG* **TAT GCA CAT GGA ACA CCT**$^{3'}$
$^{3'}$CGT GTG CC*T CGA GGG CCC ATA* **CGT GTA CCT TGT GGA**$^{5'}$

(base No 70 to 106 in pJS752-3).

The plain capital letters denote DNA originating from *elt*B, italic letters from the synthetic linker and letters in bold indicate sequences derived from the *ctx*B gene.

The sequence of the *agctb* gene is shown in Figure 25.

A.1.4. Generation of the pJS752-3 expression vector.
The *tac*P promotor is a hybrid *trp-lac* promotor that is regulated by the *lac* repressor [18]. The pJS162 plasmid carries the *lacI*^{q} gene which makes the production of rCTB under the control of the *tac*P promotor dependent of the addition of IPTG. To make the rCTB production constitutive rather than inducible the complete *agctb* gene was cut out from plasmid pJS162 with the restriction enzymes *Eco*RI and *Hind*III and introduced into the expression vector pKK223-3 (Pharmacia Biotech) digested with the same enzymes to generate pJS752-3 (Figures 18- 20). The pKK223-3 expression vector contains the *tac*P promotor but lacks the *lacI*^{q} gene. The origin of replication is derived from pBR322. Immediately downstream of the *tac* promotor is the pUC8

multiple cloning site followed by the strong transcriptional terminators T1 and T2 derived from the *rrnB* operon from *E. coli*. [19]. It has been demonstrated previously that the cloning of strong promotor signals from different sources might require the downstream placement of comparably strong termination signals for optimal transcription rates [20]. The plasmid pJS752-3 was transformed into *E.coli* HB101 (ATCC 33694).

B. The production strain *V.cholerae* 213.

B.1. The *V.cholerae* JS1569 recipient strain.

The *V.cholerae* strain JS1569 [14] is a classical O1 cholera strain originally derived from *V.cholerae* strain 569B; ATCC No 25870. The two copies of the cholera enterotoxin genes have been replaced in this strain by attenuated *ctx* genes from *V. cholerae* strain 395 that have been modified by site-directed mutagenesis [15, 21]. The attenuation comprises of a deletion of the cholera toxin A subunit gene.

The *V. cholerae* strain 395 *ctx* genes were cloned into the plasmid pCVD15 as a *Hind*III fragment. A 550 base-pair *Xba*I-*Cla*I restriction fragment was excised, deleting the A_1 gene at the position encoding the 10th amino acid and further into the gene A_2 encoding the first amino acid of the A_2 peptide. The resulting plasmid pCVD30 was digested with *Hind*III and the *Hind*III fragment cloned in the mobilisation vector pJBK85 for mobilisation into the *V. cholerae* strain 569B yielding strain CVD 103 [21]. In all 94% of the *ctx*A gene was deleted [15, 21 and SBL Vaccin test method PT00025 which method is a Polymerase Chain Reaction (PCR) with primers designed so that the deletion in the ctxA gene can be detected in both ctxA loci simultaneously (a multiplex PCR)]. To render this strain suitable for conjugation experiments a rifampicin resistant derivative of CVD103 has been selected, the JS1569 strain [14].

B.2 Introduction of the pJS752-3 plasmid into *V.cholerae* JS1569.

The pJS752-3 plasmid was purified from *E.coli* HB101 and transformed into *E.coli* S17-1 (*thi pro hsd*R *hsd*M^+ *rec*A RP4-2-Tc::Mu-Km::Tn7) [10] The plasmid was then transferred to *Vibrio cholerae* strain JS1569 by conjugation. Conjugants, *V.cholerae* (rif^R) carrying the pJS752-3 plasmid (amp^R), were selected by screening for rifampicin and ampicillin resistant colonies. The resulting strain is called *V. cholerae* strain 213. DNA sequencing of the master seed lot using purified plasmid DNA as a template confirmed the established DNA sequence (Figure 25).

C. Description of the *V.cholerae* JS1569 strain carrying the pJS752-3 expression vector; strain *V.cholerae* 213.

C.1.Nucleotide sequence of the *agctb* gene and amino acid sequence of the translated polypeptide.

C.1.1. Detailed nucleotide sequence of the* agctb *gene and flanking regions.

The complete nucleotide sequence of the pJS752-3 plasmid is given in Figure 25. The sequenced region is shown in Fig 21. Primers used for sequencing are indicated by arrows. In all approx. 600 nucleotides have been sequenced on both strands starting about 70 bp upstream the *tac*P and ending about 100 bp downstream of the coding region for the *agctb* gene. This has been done on the master seed lot as well as on cells retained after harvest from six consistency lots (Figure 26). For sequencing of plasmid DNA from master seed lots CsCl purified pJS752-3 DNA was used as template. For sequencing from cells retained after harvest from the six consistency lots a 1740 bp PCR fragment was generated (Figure 21) and sequenced. Immediately after the stop codon (TAA) there is a potential transcriptional stop (Figure 22) indicating that transcription ends at this position.

An analysis of other potential open reading frames (Figure 23) indicates two alternative open reading frames (base 900-1200 in frame 3 and base 820-1200 in frame -2). Translation of these potential sequences into amino acid sequences were made and they were analysed against the Swiss protein Data base using the Wisconsin (GCG ver 8.0) software. No significant homologies with known protein sequences were detected for the ORF in frame 3 (data not shown). Furthermore no obvious transcription or translation initiation sites preceeded this open reading frame.

The ORF in frame -2 was shown to have strong homologies (>69% over 101 amino acid overlap) with a protein orfF found both in *E.coli* and *Salmonella typhimurium*. The gene encoding this protein is in both species part of the *pyr*F operon which encodes the pyrimidine biosynthetic enzyme orotidine 5'-monophosphate (OMP) decarboxylase. The *orf*F gene product has an hithero unknown function but is thought to be associated with the OMP decarboxylase since in both *E.coli* and *S.typhimurium* the putative *orf*F translational initiation site overlaps the *pyr*F termination codon [23, 24]. It is assumed that in both *E.coli* and *S.typhimurium* the *pyr*F and *orf*F genes are transcribed into the same bicistronic messanger RNA [24]. In the pSJ752-3 expression vector the promotor proximal to the *pyrF* homolog is missing and thereby the promotor region of the operon. Consequently, it can be assumed that *orfF* is not transcribed.

C.2. Amino terminal sequence of the mature recombinant protein.

The hybrid gene *agctb* (EMBL AcNo M23050) yields when transcribed and translated into protein a mature recombinant protein having short amino acid extensions in the amino terminal of the native CTB (Figure 24). The amino acid sequence of the rCTB is otherwise identical to that reported for CTB from *V.cholerae* of classical biotype.

-3 -2 -1 +1 * * -4 -3 -2 -1 +1 +2
Ala His Gly Ala Pro Gly Tyr Ala His Gly Thr Pro
GCA CAC GGA GCT *CCC GGG* **TAT GCA CAT GGA ACA CCT**

Cleavage native LTB ↑ Cleavage native CTB ↑

Since the rCTB is a hybrid of LTB and CTB with respect to their signal peptides it also contains both signal peptidase cleavage sites. In LTB the first amino acid in the mature protein is the Ala denoted +1 and in mature CTB the first amino acid is the Thr denoted +1. The two amino acids denoted with * are introduced by the synthetic linker.

Sanchez *et al* [14] found that the leader peptidases cleaved at two sites in between the two native cleavage sites for LTB and CTB respectively. They observed cleavage both between Gly (*) and Tyr (-4) and between Tyr (-4) and Ala (-3) resulting in a recombinant protein with 3 or four extra amino terminal amino acids, namely Tyr-Ala-His-Gly or Ala-His-Gly [14].

In rCTB isolated from 1000 litre production fermentors, cleavage at four more positions were observed including the two native cleavage sites for LTB and CTB However, none of these cleavages resulted in a rCTB with an amino acid sequence shorter than the native CTB sequence.

C.3. Mode of expression.

The pKK223-2 plasmid that harbours the *agctb* gene does not contain the gene for the strong repressor (*lacI*^{q}). In *V.cholerae* it is not known if there is a repressor (*lacI*) in the genome. *V.cholerae* do not ferment lactose but have a *lacZ* gene [25]. It is reasonable to assume that the eventual repressor will not be as potent as *lacI*^{q} and also that it would be present in much smaller quantities than the promotor (*tac*P) which is located on a high copy number plasmid. As a result the expression of *agctb* is in practice constitutive.

D. Stability of the expression system.

D.1. Storage stability.

The Master and Working Seed lots of the *V.cholerae* strain 213 have been stored for more than seven and five years respectively at -65°C or colder. Genetic stability for both the Master and Working Seed lots have been demonstrated: 100% of the colonies produced rCTB when grown on LB plates containing ampicillin, 100% carried the *agctb* gene and plasmid DNA prepared from cultures of these seed lots yielded the expected DNA sequence (Figure 26).

D.2. Stability in the presence of ampicillin.

In experiments designed to investigate the stability of the plasmid retention and rCTB production, the *V.cholerae* strain 213 has been grown on a shaker in Luria Bertani broth at 37°C containing 50 µg/ml of ampicillin. Every day the culture was diluted down 10.000 fold, samples were taken and spread on LB agar with ampicillin, and the ability of the colonies to produce rCTB was tested with a colony blotting technique [SBL Vaccin test method PT00020- as described, for example, under Section D.2 (stability of the expression system)] using a monoclonal antibody specific for LTB and crossreacting with CTB. After more than 100 generations (10 days of growth) 100% of the colonies retained their capacity to produce rCTB.

D.3. Production stability.

When the *V.cholerae* strain 213 is fermented in full scale production no ampicillin is used either in the main 1000 litre fermentor or in the 4 litre preculture. Ampicillin is only used in the plates when the bacteria are grown from the frozen seed lot. To investigate the plasmid retention, and to show consistency, samples were taken at break-point from 1000 litre production fermentations (Table 4). After approx. 14-15 hours of growth at 37°C in the main 1000 litre fermentor approx. 50% of the cells have retained their capability of producing rCTB. The non-producing colonies had all lost the pJS752-3 plasmid. This was evident as the same percentage of cells producing the rCTB protein had retained the *agctb* specific DNA sequence. This also excludes the possiblity to make a copy number determination of the plasmid/chromosome ratio since only 50% of the cells contain the plasmid.

D.4. Stability of the genetic construct during production fermentation.

To show the stability of the genetic construct, DNA was prepared from breakpoint harvests of *V.cholerae* strain 213. The plasmid DNA was amplified by a PCR reaction and the resulting fragment sequenced (Figure 26). The first base in the consensus sequence corresponds to base No 5638 in pJS752-3 DNA, the last base in the consensus No 578 corresponds to base 458 in pJS752-3. The sequence determined from samples taken during production fermentation show the identical DNA sequence of the *agctb* gene and flanking control DNA to that obtained from seed lots and from the published sequence [14] (Figure 26) thereby demonstrating the stability of the construct.

Table 4. Phenotypic and genotypic stability of pJS752-3 plasmid in *V. cholerae* strain 213 after growth to break point in a 1000 liter fermentor.

Procedure.

Strain *V.cholerae* 213 was grown on LB plates containing Amp (50 µg/ml), in 5 liter preculture without ampicillin and in a 1000 liter fermentor without ampicillin. After approx. 15 hours the culture was harvested and samples taken for viable counts at the same time. For detection of phenotypic expression, colonies were lifted on nitrocellulose filters and the monoclonal antibody LT39 (LT and CTB specific) was added. After washing a rabbit anti mouse Ig-ALP conjugate was added. After further incubation the filters were developed with a Nitro tetrazodium salt [SBL Vaccin test method PT00020 as described above].

For genotypic expression, colonies were lifted on nitrocellulose filters, denatured and fixed and a radioactive probe specific for the *bla* gene (located only on the pJS752-3 plasmid) was added. After hybridization and washing under stringent conditions filters were used to expose X-ray film and positive colonies were counted [SBL Vaccin test method PT00021].

Test method PT00021

This method is very like PT00020 in that colonies are lifted from an agar plate to a nitrocellulose filter. On this filter the colonies are destroyed to denature the DNA. A single radioactive stranded DNA probe encoding a sequence only found on the expression plasmid (the amp gene for strain 213), is added, hybridized and washed under stringent condtitions. In essence, colonies are identified which have retained the plasmid but instead of looking for expression as in PT00020 you look for the DNA that can be expressed (encoding CTB)

According to PT00020 (as described above) and **PT00021.**

Results.

Culture No	% positive phenotype	% positive genotype
1550	57.3	50.8
1551	ND	50.1
1552	ND	57.8
1553	61.8	49.1
1554	ND	53.1
1555	59.6	53.3
1556	50.0	47.4

The values for culture 1555 are uncertain since the number of colonies on these plates was very low (< 10). Approx. 450-600 colonies were counted in each assay from the remaining cultures.

Comparative Examples I

"401" strain vs "213" strain

213 production system

The rCTB component which is produced in a 213 *V. cholerae* expression system has already been fully described but is summarised as follows:

A *V. cholerae* 01 Inaba strain 569B of the classical biotype (ATCC No 25870) was genetically modified by removal of the gene sequence encoding the CTA gene was removed (CVD103 strain). The strain was made rifampicin resistant by insertion of a rifampicin resistance marker into the CVD103 strain to produce a JS1 569 strain.

The modified cholerae strain (JS1569) was then transfected with a plasmid (designated pJS752-3) containing the gene sequence for CTB under the control of a heterologous promoter, wherein the CTB coding sequence is linked to a sequence encoding a heterologous leader polypeptide (the *E. coli* LT leader sequence) to facilitate secretion of CTB from the host cell. The pJS752-3 plasmid was prepared by excision of the CTB gene in plasmid JS162 and inserting the gene into a plasmid vector PKK223-1 which contains the tacP promoter but not the laclq gene present in PJS162 that is responsible for IPTG dependence (more details on the methods of preparing and using these plasmids are described herein and are provided in Sanchez and Holmgren (1989) ibid and US patents Nos 5268276, 58234246 and 6043057 and EP Patent No 0368819B).

The pJS752-3 plasmid further comprises an antibiotic selectable marker (ampicillin resistance marker) to enable selection of suitable plasmids containing the CTB sequence. The designation of the *V. cholerae* production strain (JS1569) with the expression vector (pJS752-3) is the *V. cholerae* 213 strain.

The CTB was overexpressed and secreted from the 213 production strain in monomeric form, whereafter it assembles into the characteristic pentameric ring-like structure to provide rCTB having a molecular weight of approx 58kDa. In this way, the rCTB consists only of the non-toxic part of the cholera enterotoxin (since the toxic A subunit has been genetically deleted from the production strain) but retains its ability to bind GM-1 receptors on the surface of intestinal epithelial cells (see US patents Nos 5268276, 58234246 and 6043057, EP Patent No 0368819 and Sanchez *et al* (1989) (*ibid*) for more details on this expression system)

"401" Expression System

As described herein, a derivative of the JS 1569 *V. cholerae* production strain which lacks the functionality of a *thyA* gene has been produced (for example, the *thyA* gene may be removed or may be genetically disabled). A functional *thyA* gene is provided in an expression plasmid which allows for the selection of *V. cholerae* host cells which retain the plasmid and which are unable to grow in the absence of thymine (as described in WO 99/61634). The designation of the derived *V. cholerae* production strain (JS1569 Δ*thyA*Δkan) with the expression vector (pMT-CtxB*thyA*-2) as described herein is the CTB producing 401 *V. cholerae* strain.

Table 5: Comparison of yield of rCTB-401 and rCTB-213 in three consecutive fermentations

	rCTB-401 in year 1998			**rCTB-213 in year 1999**			**Average yield of rCTB-213 in 1999-2001 (about 60 batches)**
Batch nr	**M4806**	**M4807**	**M4808**	**RC2902**	**RC2903**	**RC2904**	
mg rCTB/ml	1,2	1,4	1,5	0,44	0,48	0,52	0,41±0,07 (±1 SD)

The data in Table 5 demonstrates that the average yield of rCTB of about 1.4mg/ml from the *thyA* deleted *V. cholerae* strain (termed the "401" strain) at the end of the fermentation period is in the range of 3-4 times greater than the yield of rCTB (0.4mg/ml) from the "213" strain).

The method used for measuring the rCTB concentration is single radial immunodiffusion (SRI) also known as the Mancini test (Mancini *et al* (1965) Immunochem 2: 235-254: Immunochemical quantitation of antigen by single radial immunodiffusion) using antisera against highly purified rCTB. The rCTB standard used for preparing a calibration curve is a highly purified rCTB which has been characterised by a number of protein tests.

This yield (1.4mg/ml) is about 50 fold higher than that reported from wild type *V. cholerae* 569B strain and about 20-fold higher than that reported by Sanchez and Holmgren (1989).

Discussion of Comparative Examples I

The main differences between the expression plasmids used in the "213" production system and the "401" production system are outlined in Table 3. These include the absence of an antibiotic resistance marker, a smaller plasmid size and a higher yield of rCTB from the "401" strain relative to the "213" strain.

Without wishing to be bound by theory, it is believed that removing a portion of non-coding *V. cholerae* DNA downstream of the *ctxB* gene resulting in the reduced size of the expression vector contributes to the improved stability and the improved yield of the CTB end product. By way of example of the improved plasmid stability, the plasmid containing the cassette was still present in 100% of the bacterial cells in a culture after 100 generations even in the absence of antibiotic selection. The absence of an antibiotic resistance marker in the "401" strain also has advantages in terms of a safer and cheaper CTB end product. The produced rCTB is also advantageous because a more homogeneous CTB product is produced, In this respect, when the production strain is the 401 strain, only one rCTB is produced and this rCTB sequence differs from the wild type CTB sequence only in a single N terminal mutation (substitution of Threonine (Thr) to Alanine (Ala)). In constrast, when the production strain is the 213 strain, the final rCTB product actually contains slightly different rCTB amino acid sequences (see Sanchez and Holmgren 1989 (*ibid*)) as there are at least two different mutations occurring within the N-terminal residues of the CTB sequence.

Comparative Examples II

Levels of CTB produced using 358 strain (Lebens) and 401 expression system

The main differences between the expression plasmids used in the Lebens *et al* (1993) (*ibid*) production system and the "401" production system are outlined in Table 3. These include the absence of an antibiotic resistance marker and a smaller plasmid size.

The CTB expression system described in Lebens *et al* (1993) requires the presence of an antibiotic whenever the organism is grown. The antibiotic resistance marker is an ampicillin resistance marker. The ampicillin resistance is due to the expression of the enzyme b-lactamase which cleaves the antibiotic. The *V. cholerae* strain termed the "358" strain used in the Lebens expression system requires the continuous presence of ampicillin in the medium in order to maintain optimum production. Thus, the CTB yields obtained using the Lebens expression system are only obtained using "selective pressure and in the presence of ampicillin".

Comparative Example II

Levels of CTB produced using the expression system disclosed in WO 01/27144 and the "401" expression system

Production of Recombinant CTB in Bacteria

The expression plasmid MS-0 (see Figure 2 of WO 01/27144) was used to express rCTB and variants thereof. MS-0 containing the rCTB gene is named pML-CTBtacI. The plasmid pML-CTBtacI surprisingly generates up to five times the product which was generated by a comparable plasmid (Vector pJS162 as disclosed in Sanchez and Holmgren 1989 *ibid*). pML-CTBtacI was constructed by cloning a portion of the CTB genomic region and the complete CTB coding region into plasmid MS-0 creating a 3.66 Kb expression plasmid. The PvuII site in the polylinker was destroyed during cloning. The plasmid contains a tac promoter from pKK223, an EcoRI-BamHI polylinker fragment, and can be found at genbank accession No M77749.

The encoded protein is identical to the sequence from V. cholera strain 569B (SEQ ID NO: 2). The signal sequence (SEQ ID NO: 3) is also from the CTB V. cholera classic strain 569B CTB gene. The complete nucleotide sequence of V. cholera strain 569B CTB gene is shown in Figure 1 (SEQ ID NO: 1) of WO 01/27144. The signal sequence for LTB (SEQ ID NO : 15) is MNKVKCYVLFTALLSSLCAYG is also shown in the sequence listing of WO 01/27144 and can be used in the production of mutants or variants of LTB.

Comparison with the 401 expression system

The main differences between the expression plasmid used in the CTB production system disclosed in WO 01/27144 and the "401" production system are outlined in Table 3. These include the absence of an antibiotic resistance marker, a smaller plasmid size and a higher yield of rCTB from the "401" strain relative to the yields obtained from the expression system disclosed in WO 01/27144.

Overall Summary

It is well known that high level transcription and translation of proteins depends on many factors. These factors include but are not limited to: promoter strength, translational initiation sequences, codon choice, secondary structure of mRNA, transcriptional termination, plasmid copy number, plasmid stability and host cell physiology. Thus, the expression of different proteins can vary dramatically and the use of a strong promoter alone does not guarantee the successful overexpression of a desired protein.

This present invention teaches how to improve CTB yields using a CTB production system with markers other than antibiotic resistance markers and appropriate host cell strains that remove the need for antibiotic selection. The CTB production system comprises a bacterial host cell lacking the functionality of a *thyA* gene which is used in conjunction with a novel expression vector to produce unexpected high yields of recombinant B subunit of the cholera toxin (rCTB) relative to the yields obtained with known bacterial host cell production systems.

In one embodiment, the present invention teaches how to improve CTB yields using a CTB production system comprising a *V. cholerae* host cell lacking the functionality of a *thyA* gene which is used in conjunction with a novel expression vector to produce unexpected high yields of recombinant B subunit of the cholera toxin (rCTB) relative to the yields obtained with known *V. cholerae* production systems.

A plasmid expression vector was constructed in which the thymidylate synthase (*thyA*) gene of *E. coli* was used as a means of selection and maintenance of a plasmid comprising a CTB gene. The plasmids is of reduced size relative to known expression plasmids for producing CTB because substantially all of the non coding *V. cholerae* DNA downstream of the CTB gene was removed.

The unexpected high yield of CTB obtained using this expression system demonstrated both the efficiency of expression of heterologous genes in the *V. cholerae* strain and the stability of the plasmids maintained by complementation of the *thyA* deletion. Furthermore, the plasmid was found to be extremely stable. Even after repeated passages through liquid culture equivalent to 100 generations all the cells retained the plasmid and the ability to express the recombinant protein.

The expression system as reported here is advantageous because it facilitates the production of CTB for the following uses which include but are not limited to:
a protective immunogen in oral vaccination against cholera and LT-caused *E. coli* diarrhoea;
An immunomodulator or a tolerogenic inducing agent or an immune-deviating agent for down-regulating/modulating/de-sensitising/re-directing the immune response;
An adjuvant for altering, enhancing, directing, re-directing, potentiating or initiating an antigen-specific or non-specific immune response;
A carrier to stimulating an immune response to one or more unrelated antigens; and
A diagnostic agent for producing antibodies (such as monoclonal or polyclonal antibodies) for use in diagnostic or immunodiagnostic tests.

It is a particular advantage from the point of purification and standardisation of CTB as a vaccine component that relatively high yields of CTB can be achieved using stable bacterial host cell strains that lack the functionality of a *thyA* gene.

The present invention also teaches how to obtain a stable CTB preparation which is essentially free of antibiotic residues resulting in a safer product for human use.

Spirit and Scope of the Invention

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that all values are approximate, and are provided for description. Patents, patent applications, publications, product descriptions, and protocols are cited throughout this application, the disclosures of which are incorporated herein by reference in their entireties for all purposes. In the case of inconsistencies, the present disclosure, including definitions, will prevail.

References

1.Lebens, M., and J. Holmgren. 1994. Structure and arrangement of the cholera toxin genes in *V. cholerae* O139. FEMS Microbiology Letters. **117**:197-202.

2. Kaper, J. B., H. Lockman, M. M. Baldini, and M. M. Levine. 1984. A recombinant live oral cholera vaccine. Biotechnology **2**:345-349.

3.Sandkvist, M., T. R. Hirst, and M. Bagdasarian. 1987. Alterations at the carboxyl terminus change assembly and secretion properties of the B subunit of *Escherichia coli* heat-labile enterotoxin. J. Bacteriol. **169**:4570-4576.

4.Lebens, M., S. Johansson, J. Osek, M. Lindblad and J. Holmgren. 1993. Large scale production of *V. cholerae* toxin B subunit for use in oral vaccines. Biotechnology **11**:1574-1578.

5.Miller, V. L., and J. J. Mekalanos. 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *V. cholerae* requires *toxR* J. Bacteriol. **170**:2575-2583.

6.Milton, D. L., A. Nordqvist and H. Wolf-Watz. 1992. Cloning of a Metalloprotease gene involved in the virulence mechanism of *Vibrio anguillarum.* J. Bacteriol. **174**:7235-7244.

7. Milton, D. L., R. O'Toole, P. Hörstedt and H. Wolf-Watz. 1996. Flagellin A is essential for the virulence of *Vibrio anguillarium.* J. Bacteriol. **178**:1310-1319.

8. Levine, M. M., J.B. Kaper, D. Herrington, J. Ketley, G. Losonsky, C.O. Tacket, B. Tall and S. Cryz. 1988. Safety, immunogenicity, and efficacy of recombinant live oral cholera vaccines, CVD 103 and CVD 103-HgR. Lancet *ii*:467-470.

9. Sanchez, J., and J. Holmgren. 1989. Recombinant system for overexpression of cholera toxin B subunit in *V. cholerae* as a basis for vaccine development. Proc. Natl. Acad. Sci. USA. **86**:481-485.

12. Parsot, C. 1992. Identification of a *lacZ* gene in *V. cholerae*. Res. Microbiol. **143**:27-36.

13.Lebens, M., and J. Holmgren. 1994. Structure and arrangement of the cholera toxin genes in *V. cholerae* O139. FEMS Microbiology Letters. **117**:197-202.

14.Sanchez, J., and J. Holmgren. 1989. Recombinant system for overexpression of cholera toxin B subunit in *V. cholerae* as a basis for vaccine development. Proc. Natl. Acad. Sci. USA. **86**:481-485.

15.Kaper, J. B., H. Lockman, M. M. Baldini, and M. M. Levine. 1984. A recombinant live oral cholera vaccine. Biotechnology **2**:345-349.

16.Sandkvist, M., T. R. Hirst, and M. Bagdasarian. 1987. Alterations at the carboxyl terminus change assembly and secretion properties of the B subunit of *Escherichia coli* heat-labile enterotoxin. J. Bacteriol. **169**:4570-4576.

17.Fürste, J. P., W. Pansegrau, R. Frank, H. Blöcker, P. Scholtz, M. Bagdasarian, and E. Lanka. 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tac*P expression vector. Gene **48**:119-131.

18. Sambrook, J., E. F. Fritsch, amd T. Maniatis. 1989. Molecular cloning. A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

19.Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller. 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli.* J. Mol. Biol. 148:107-127.

20.Gentz, R., A. Langner, A. C. Y. Chang, S. N. Cohen, and H. Bujard. 1981. Cloning and analysis of strong promotors is made possible by the downstream placement of a RNA termination signal. Proc. Natl. Acad. Sci. USA, 78:4936-4940.

21. Levine, M. M., J.B. Kaper, D. Herrington, J. Ketley, G. Losonsky, C.O. Tacket, B. Tall and S. Cryz. 1988. Safety, immunogenicity, and efficacy of recombinant live oral cholera vaccines, CVD 103 and CVD 103-HgR. Lancet *ii*:467-470.

22. Simon, R., U. Priefer, and A. Pühler. 1983. A broad host range mobilization system for *in vivo* genetic engineering; transposon mutagenesis in Gram negative bacteria. Biotechnology **1**:784-791.

23.Turnbough, C. L., K. H. Kerr, W. R. Funderberg, J. P. Donahue and F. E. Powell. 1987. Nucelotide sequence and characterization of the *pyr*F operon of *Escherichia coli* K12. J. Biol. Chem. **262**:10239-10245.

24.Theisen, M., R. A. Kelln and J. Neuhard. 1987. Cloning and characterization of the *pyr*F operon of *Salmonella typhimurium.* Eur. J. Biochem. **164**:613-619.

25.Parsot, C. 1992. Identification of a *lac*Z gene in *V. cholerae*. Res. Microbiol. 143:27-36.

CLAIMS

1. An expression system for producing a B subunit of a cholera toxin (CTB) wherein the expression system comprises:
   (a) a bacterial host cell lacking the functionality of a *thyA* gene; and
   (b) an expression vector comprising a functional *thyA* gene and a CTB gene which is substantially free of the flanking sequences immediately contiguous by the 5' and 3' end of the CTB gene in the naturally occurring genome of the host cell from which the CTB gene is derived.

2. An expression system according to claim 1 wherein the host cell is from the genus *Vibrio.*

3. An expression system according to claim 1 or claim 2 wherein the host cell is a *V. cholerae* host cell.

4. An expression system according to any one of claims 1-3 wherein the host cell is an attenuated host cell.

5. An expression system according to claim 4 wherein the host cell lacks the functionality of a CTA gene.

6. An expression system according to any one of claims 1-5 wherein the expression vector is less than 5kb in size, preferably about 3kb in size.

7. An expression system according to any one of claims 1-6 wherein the CTB gene is operably linked to a heterologous promoter.

8. An expression system according to claim 7 wherein the expression vector further comprises a nucleotide sequence encoding a leader polypeptide operably linked to the CTB gene and the heterologous promoter.

9. An expression system according to claim 7 or claim 8 herein the heterologous promoter is a tacP promoter.

10. An expression system according to claim 8 or claim 9 wherein the leader polypeptide is the leader polypeptide of *E. coli* heat labile enterotoxin.

11. An expression system according to claim 10 wherein the leader polypeptide is the leader polypeptide of the B subunit of *E. coli* heat labile enterotoxin (LTB).

12. An expression system according to any one of the previous claims wherein the expression vector comprises an *E. coli thyA* gene.

13. An expression system according to any one of the previous claims wherein the expression vector has the sequence presented in Figure 16.

14. An expression system according to any one of claims 1-13 wherein the expression vector further comprises at least one further nucleotide sequence encoding a heterologous protein.

15. An expression system according to claim 14 wherein the further nucleotide sequence encodes a non-toxic component or form of the heat labile *E. coli* enterotoxin LT, preferably the non-toxic component of LT is the B subunit of a (LTB) or a fragment thereof.

16. A method of producing CTB wherein the method comprises:
transforming a bacterial host cell according to any one of claims 1-5 with an expression vector as defined in any one of claims 6-15; and
culturing the transformed bacterial host cell under conditions which permit production of the CTB.

17. The method of claim 16 wherein the method further comprises isolating and/or purifying the CTB from the host cell.

18. An isolated nucleic acid construct which comprises a *thyA* gene and a CTB gene which is substantially free of the flanking sequences immediately contiguous by the 5' and 3' end of the CTB gene in the naturally occurring genome of the host cell from which the CTB gene is derived.

19. A nucleic acid construct according to claim 18 wherein the nucleic acid construct is less than 5kb in size, preferably about 3kb in size.

20. A nucleic acid construct according to claim 18 or 19 wherein the CTB gene is operably linked to a heterologous promoter.

21. A nucleic acid construct according to claim 20 wherein the nucleic acid construct further comprises a nucleotide sequence encoding a leader polypeptide operably linked to the CTB gene and the heterologous promoter.

22. A nucleic acid construct according to claim 20 or 21 wherein the heterologous promoter is a tacP promoter.

23. A nucleic acid construct according to claim 21 or 22 wherein the leader polypeptide is the leader polypeptide of *E. coli* heat labile enterotoxin.

24. A nucleic acid construct according to claim 23 wherein the leader polypeptide is the leader polypeptide of the B subunit of *E. coli* heat labile enterotoxin (LTB).

25. A vector comprising the nucleic acid construct according to any one of claims 18-24.

26. A vector according to claim 25 wherein the vector is a plasmid.

27. A host cell comprising the vector according to claim 25 or claim 26.

28. A host cell according to claim 27 wherein the host cell is a *V. cholerae* host cell.

29. A *V. cholerae* host cell according to claim 28 wherein the host cell does not express an antibiotic resistance marker.

30. A host cell according to claim 28 or claim 29 wherein the host cell is attenuated by deletion or inactivation of one or more genes.

31. An isolated plasmid pMT-ctxB*thyA*-2 characterised by a restriction endonuclease map as shown in Figure 14.

32. An isolated nucleotide sequence as shown in Figure 16.

33. An isolated CTB preparation wherein the CTB preparation is essentially free of antibiotic residues.

34. An isolated expressed CTB according to claim 33 or claim 34 wherein the CTB has the N-terminal amino acid sequence APQNIT (Ala-Pro-Gln-Asn-Ile).

35. A vaccine against an enterotoxin-induced illness wherein the vaccine comprises CTB according to claim 33 or 34 and a pharmaceutically acceptable carrier or diluent.

36. A vaccine according to claim 35 wherein the vaccine further comprises a whole cell antigen.

37. A CTB according to claim 33 or 34 or a vaccine according to claim 35 or claim 36 for use in medicine.

38. Use of CTB according to claim 33 or 34 or a vaccine according to claim 35 or claim 36 in the preparation of a medicament for the prevention and/or treatment of an enterotoxin-induced illness.

39. A method of preventing and/or treating an enterotoxin-induced illness in a host mammalian subject wherein the method comprises administering to the mammal a CTB according to claim 33 or 34 or a vaccine according to claim 35 or claim 36.

40. Use of a transdermally administred composition comprising CTB in the preparation of a medicament for the prevention and/or treatment of an enterotoxin-induced illness.

41. Use according to claim 40 wherein composition further comprises one or more antigens.

42. Use according to claim 41 wherein the antigen is a whole cell antigen.

43. Use according to any one of claims 40-41 wherein the CTB has the N-terminal amino acid sequence APQNIT.

44. Use according to any one of claims 40-43 wherein the CTB is produced using the expression system of any one of claims 1-15 or by the method of claim 16 or 17.

45. An inactivated *V. cholerae* whole cell preparation comprising a CTB molecule essentially free of antibiotic residues.

46. An inactivated *V. cholerae* whole cell preparation according to claim 45 for use in medicine.

47. Use of the Vibrio cholerae preparation according to claim 45 in the manufacture of a medicament for the prevention and/or treatment of an enterotoxin-induced illness.

48. An expression system substantially as described herein and with reference to the accompanying Figures.

49. An isolated nucleic acid construct substantially as described herein and with reference to the accompanying Figures.

50. An isolated stably expressed CTB preparation substantially as described herein and with reference to the accompanying Figures.

ABSTRACT
EXPRESSION SYSTEM

The present invention provides an expression system for producing a B subunit of a cholera toxin (CTB) wherein the expression system comprises a bacterial host cell lacking the functionality of a *thyA* gene; and an expression vector comprising a functional *thyA* gene and a substantially pure CTB gene which is substantially free of the flanking sequences immediately contiguous by the 5' and 3' end of the CTB gene in the naturally occurring genome of the host cell from which the CTB gene is derived. The present invention also provides an expression vector for use in the expression system, an isolated CTB preparation which is essentially free of antibiotic residues, vaccines comprising the CTB preparation and uses thereof.

Figure 1

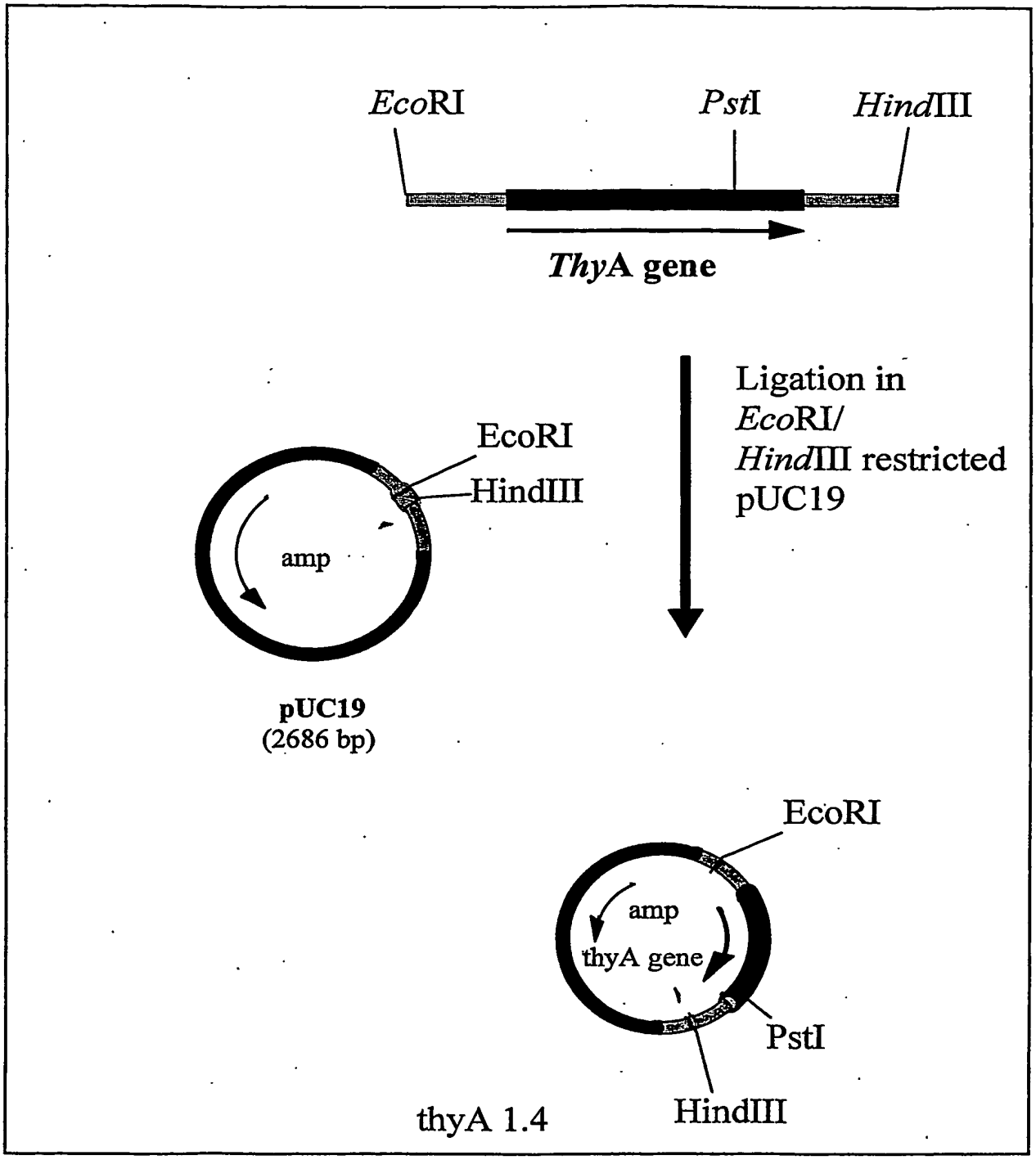

*Eco*RI
*Pst*I
*Hind*III
*Thy*A gene
Ligation in
*Eco*RI/
*Hind*III restricted
pUC19
EcoRI
HindIII
amp
pUC19
(2686 bp)
EcoRI
amp
thyA gene
PstI
HindIII
thyA 1.4

Figure 2

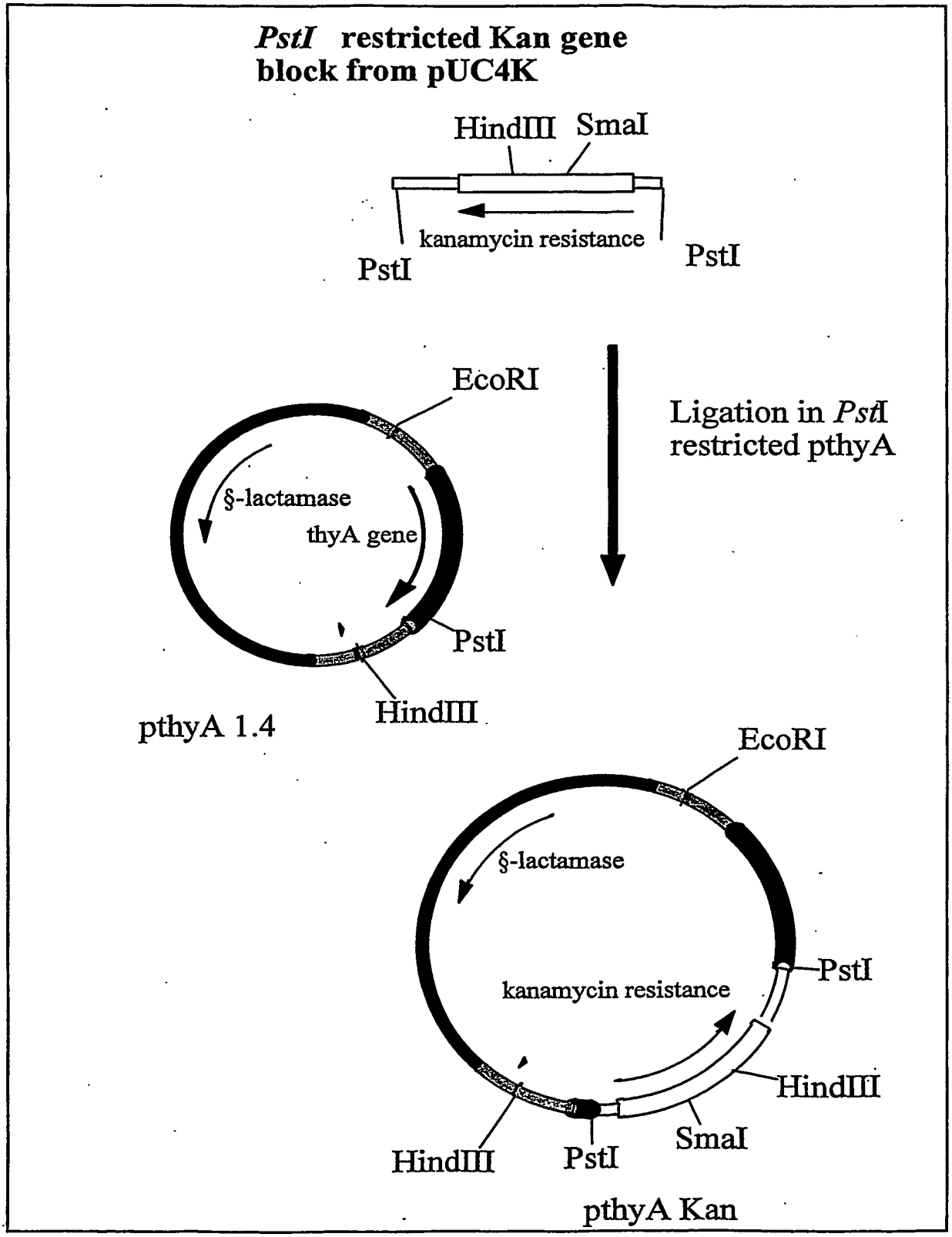
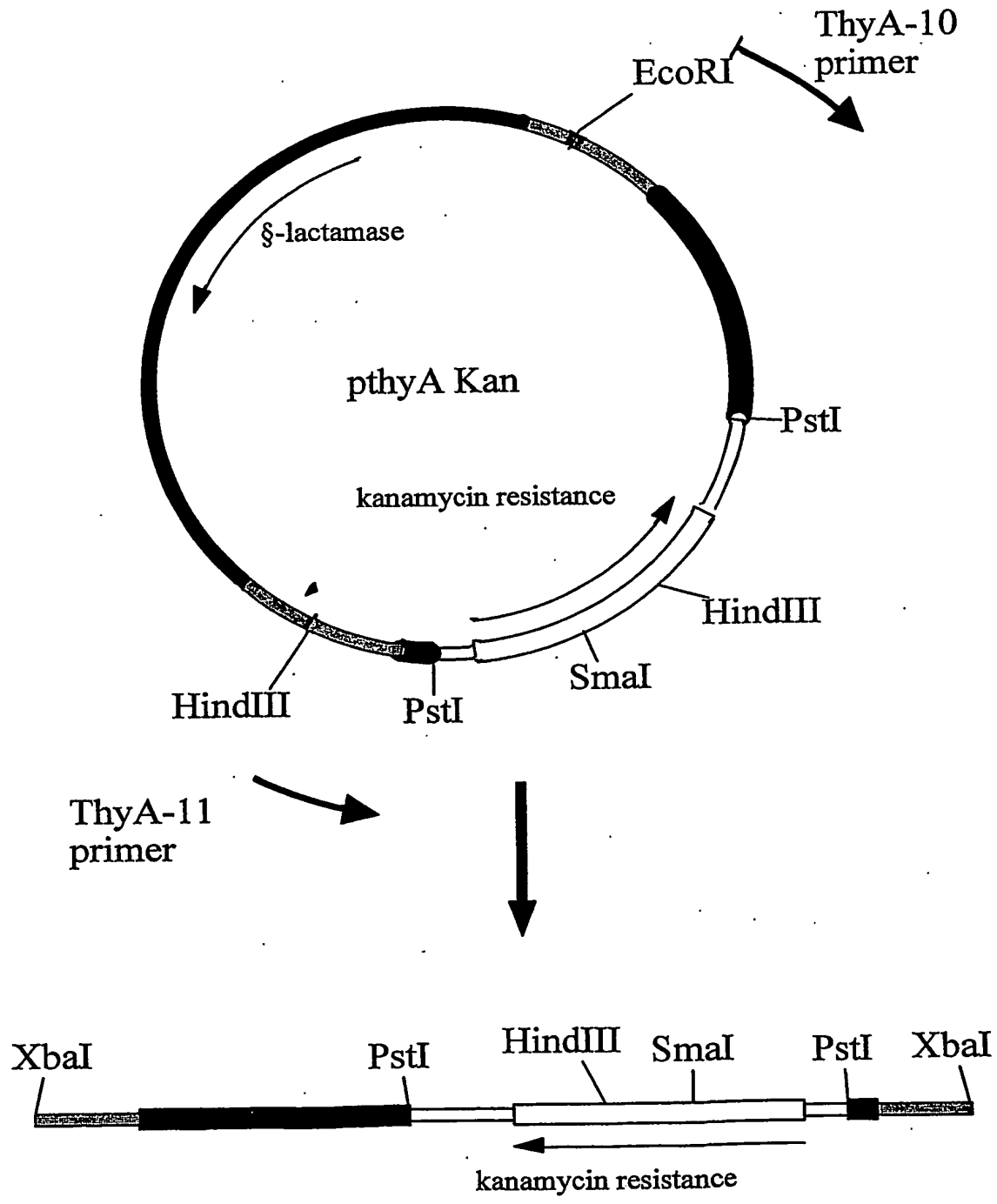

PstI restricted Kan gene block from pUC4K
HindIII
SmaI
kanamycin resistance
PstI
PstI
EcoRI
Ligation in *PstI* restricted pthyA
§-lactamase
thyA gene
PstI
HindIII
pthyA 1.4
EcoRI
§-lactamase
PstI
kanamycin resistance
HindIII
SmaI
HindIII
PstI
pthyA Kan

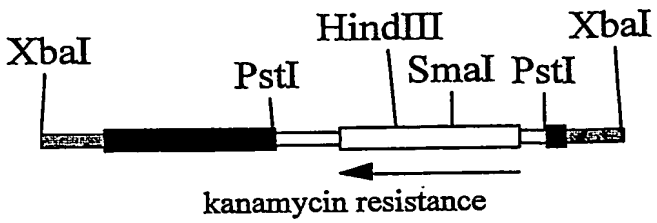
Figure 3

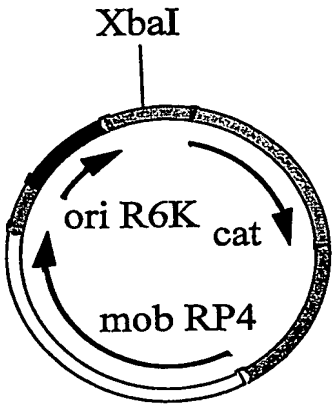
PCR to generate thyA-Kan fragment with *Xba*I ends. Primers were choosen so that the *Eco*RI and *Hind*III sites were eliminated



thyA Kan XbaI ends

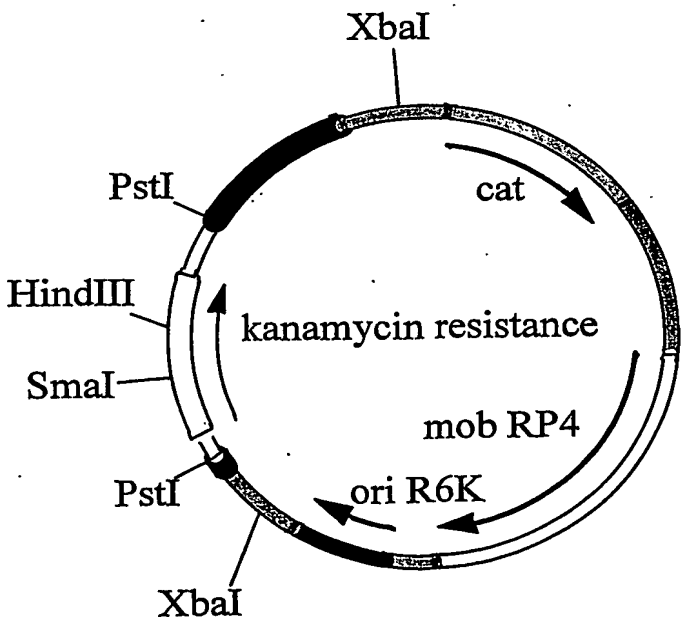
Figure 4





pNQ 705
(4450 bp)

Ligation of *thy*A-kan fragment with *Xba*I ends in *Xba*I restricted pNQ705



pNQ705 thyA KanR

Figure 5

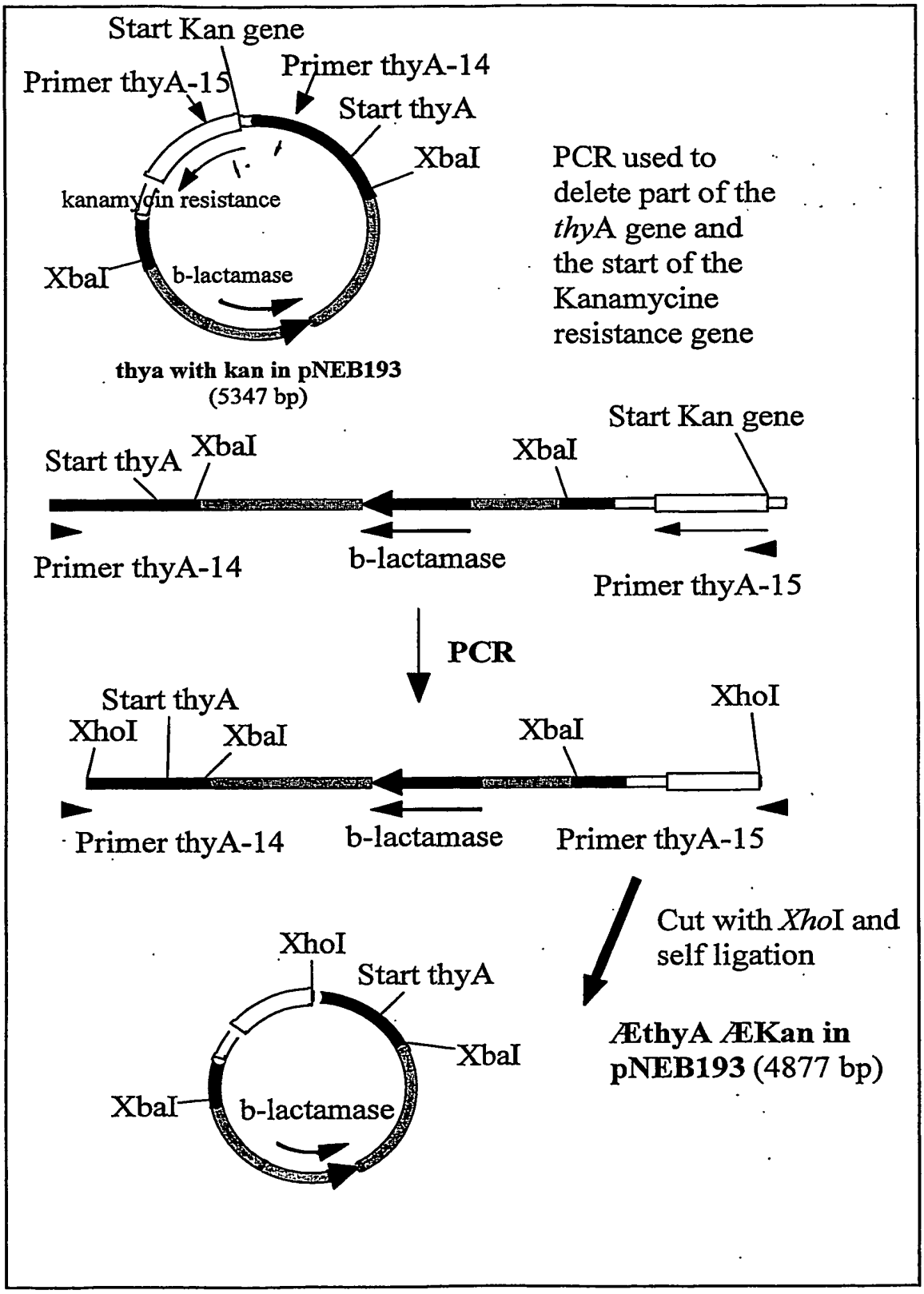

Start Kan gene
Primer thyA-15
Primer thyA-14
Start thyA
XbaI
kanamycin resistance
XbaI
b-lactamase
thya with kan in pNEB193
(5347 bp)
PCR used to delete part of the thyA gene and the start of the Kanamycine resistance gene
Start thyA
XbaI
XbaI
Start Kan gene
Primer thyA-14
b-lactamase
Primer thyA-15
PCR
Start thyA
XhoI
XbaI
XbaI
XhoI
Primer thyA-14
b-lactamase
Primer thyA-15
Cut with XhoI and self ligation
XhoI
Start thyA
XbaI
XbaI
b-lactamase
ÆthyA ÆKan in pNEB193 (4877 bp)

Figure 6

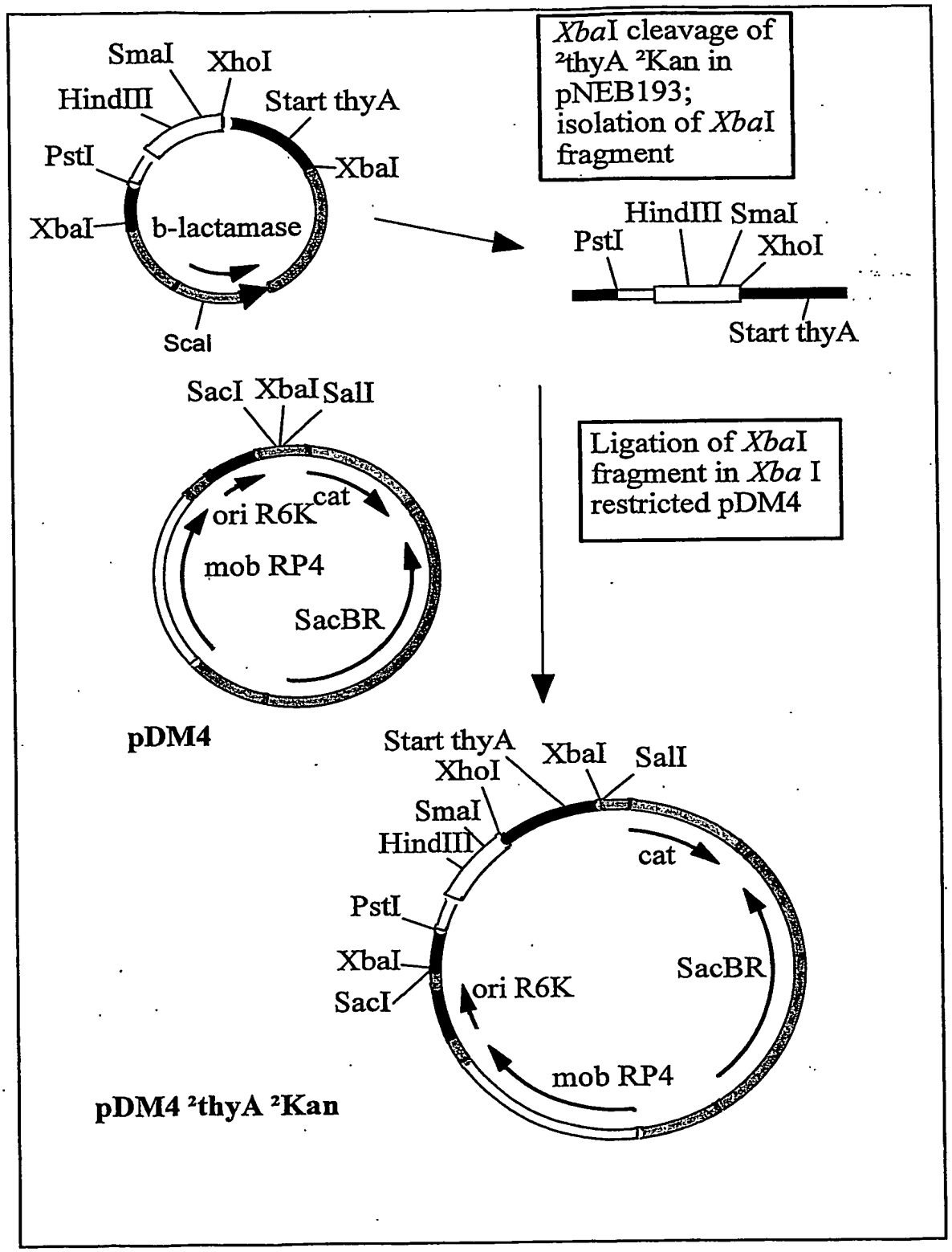

SmaI
XhoI
HindIII
Start thyA
PstI
XbaI
XbaI
b-lactamase
ScaI
XbaI cleavage of ²thyA ²Kan in pNEB193; isolation of XbaI fragment
HindIII SmaI
PstI
XhoI
Start thyA
SacI XbaI SalI
ori R6K
cat
mob RP4
SacBR
pDM4
Ligation of XbaI fragment in Xba I restricted pDM4
Start thyA
XhoI
XbaI
SalI
SmaI
HindIII
cat
PstI
XbaI
SacI
ori R6K
SacBR
mob RP4
pDM4 ²thyA ²Kan

Figure 7

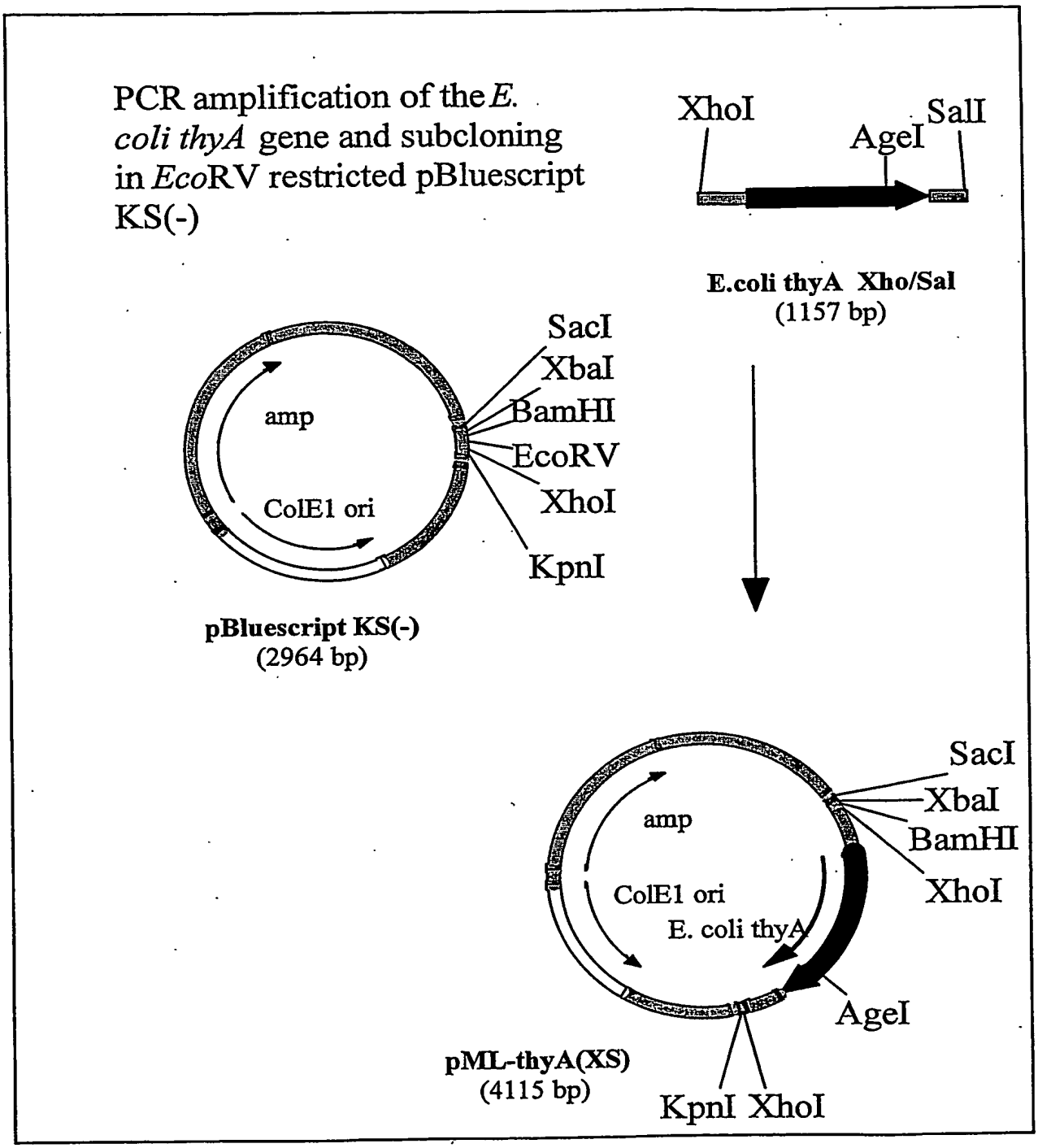

PCR amplification of the *E. coli thyA* gene and subcloning in *Eco*RV restricted pBluescript KS(-)
XhoI
AgeI
SalI
E.coli thyA Xho/Sal
(1157 bp)
SacI
XbaI
BamHI
EcoRV
XhoI
KpnI
amp
ColE1 ori
pBluescript KS(-)
(2964 bp)
SacI
XbaI
BamHI
XhoI
amp
ColE1 ori
E. coli thyA
AgeI
pML-thyA(XS)
(4115 bp)
KpnI XhoI

Figure 8

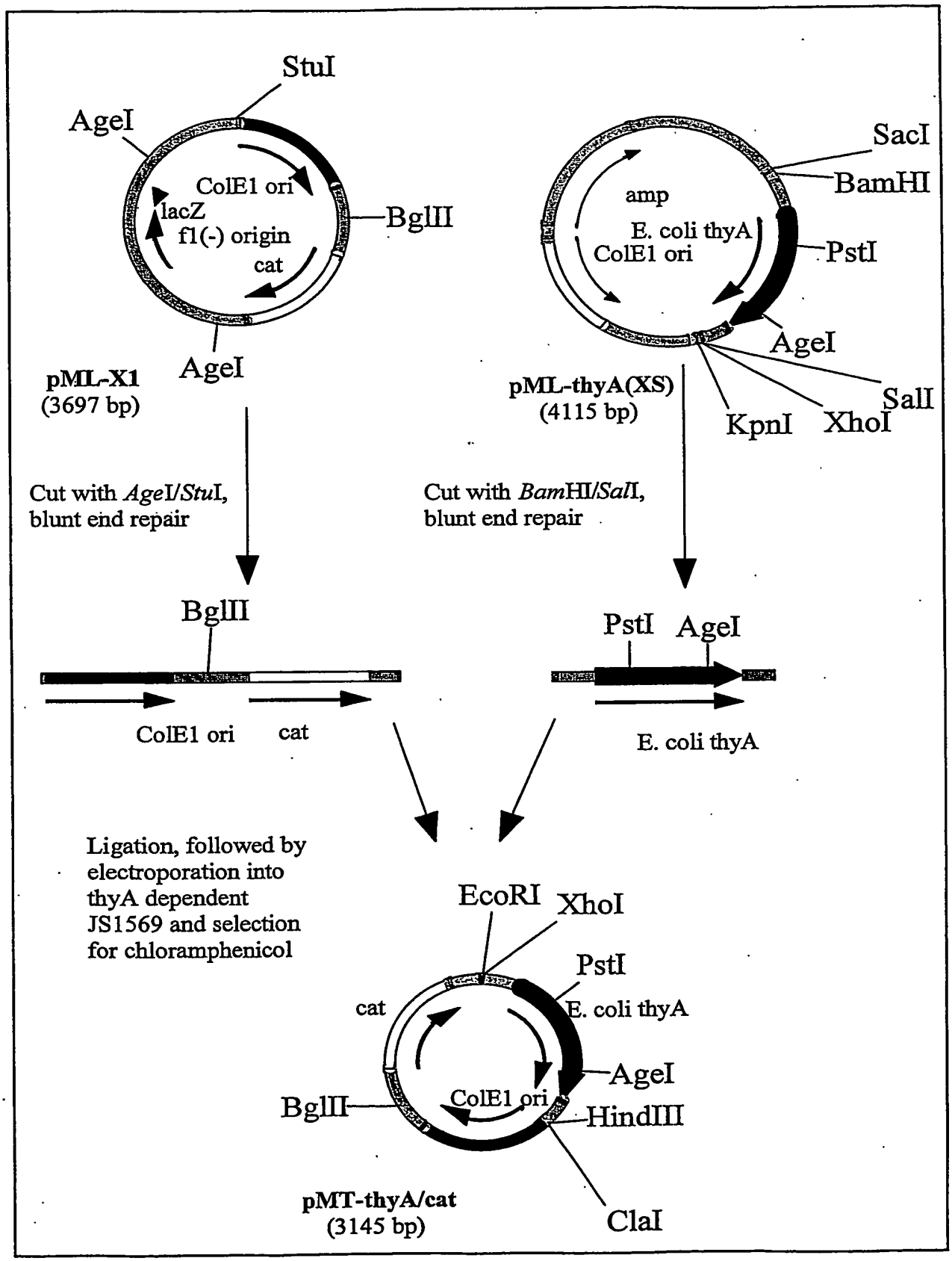

StuI
AgeI
ColE1 ori
lacZ
f1(-) origin
cat
BglII
AgeI
pML-X1
(3697 bp)
SacI
BamHI
amp
E. coli thyA
ColE1 ori
PstI
AgeI
SalI
KpnI
XhoI
pML-thyA(XS)
(4115 bp)
Cut with AgeI/StuI, blunt end repair
Cut with BamHI/SalI, blunt end repair
BglII
ColE1 ori
cat
PstI
AgeI
E. coli thyA
Ligation, followed by electroporation into thyA dependent JS1569 and selection for chloramphenicol
EcoRI
XhoI
PstI
cat
E. coli thyA
AgeI
BglII
ColE1 ori
HindIII
ClaI
pMT-thyA/cat
(3145 bp)

Figure 9

*Sac*I

TCC TCT CTA TGT GCA CAC G*GA GCT C*CT CAA AAT ATT — Sequence in pMTctxB (thyA)-2

Ser Ser Leu Cys Ala His Gly **Ala** Pro Gln Asn Ile

TCC TCA GCA TAT GCA CAT GGA ACA CCT CAA AAT ATT — Sequence in native *ctxB*

Ser Ser Ala Tyr Ala His Gly **Thr** Pro Gln Asn Ile

TCC TCT CTA TGT GCA CAC GGA GCT CCT CAG TCT ATT — Sequence in native *eltB*

Ser Ser Leu Cys Ala His Gly Ala Pro Gln Ser Ile

Figure 10

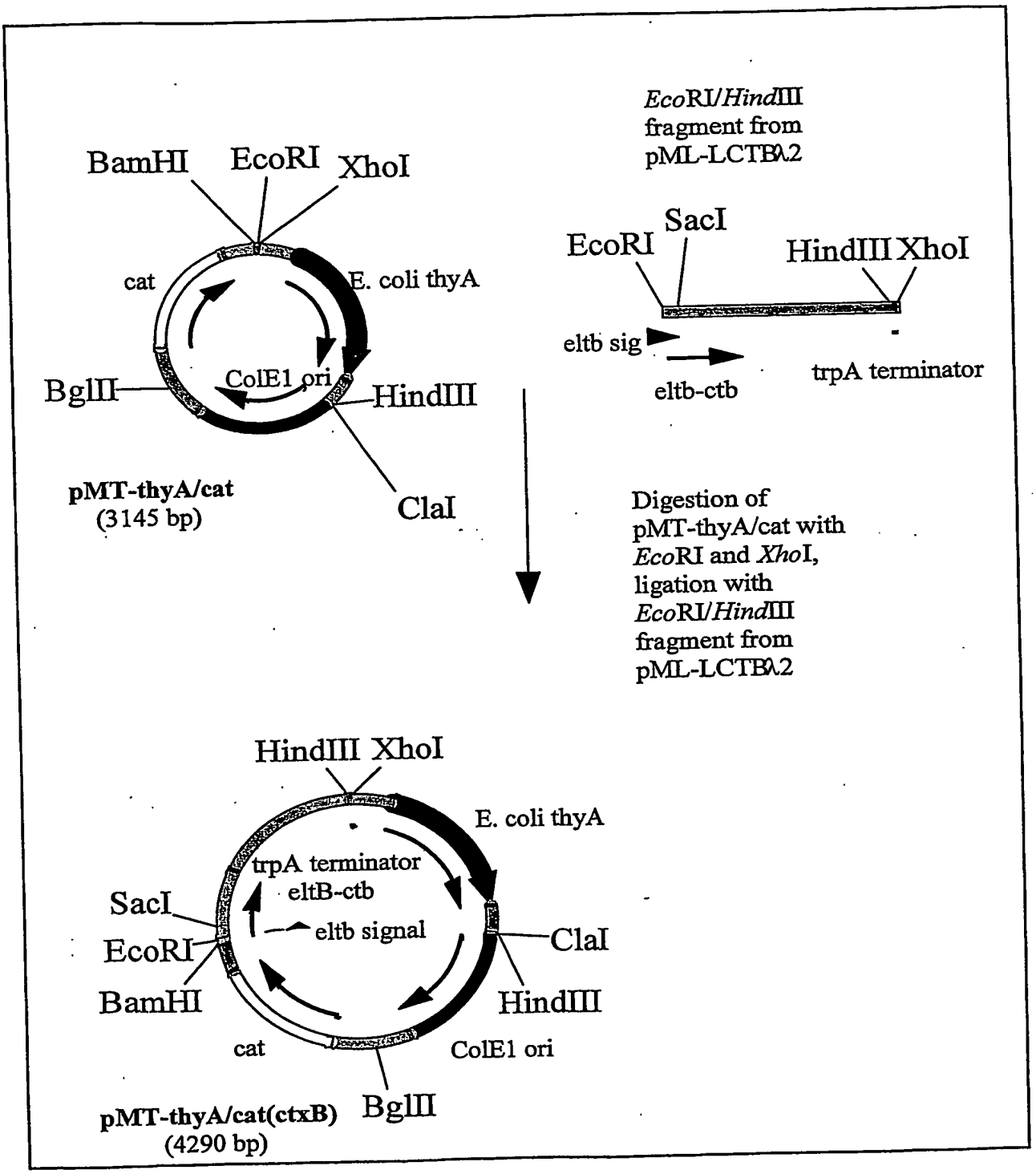

BamHI
EcoRI
XhoI
cat
E. coli thyA
BglII
ColE1 ori
HindIII
ClaI
pMT-thyA/cat
(3145 bp)
EcoRI/HindIII
fragment from
pML-LCTBλ2
EcoRI
SacI
HindIII
XhoI
eltb sig
eltb-ctb
trpA terminator
Digestion of
pMT-thyA/cat with
EcoRI and XhoI,
ligation with
EcoRI/HindIII
fragment from
pML-LCTBλ2
HindIII
XhoI
E. coli thyA
trpA terminator
eltB-ctb
eltb signal
SacI
EcoRI
BamHI
ClaI
HindIII
cat
ColE1 ori
BglII
pMT-thyA/cat(ctxB)
(4290 bp)

Figure 11

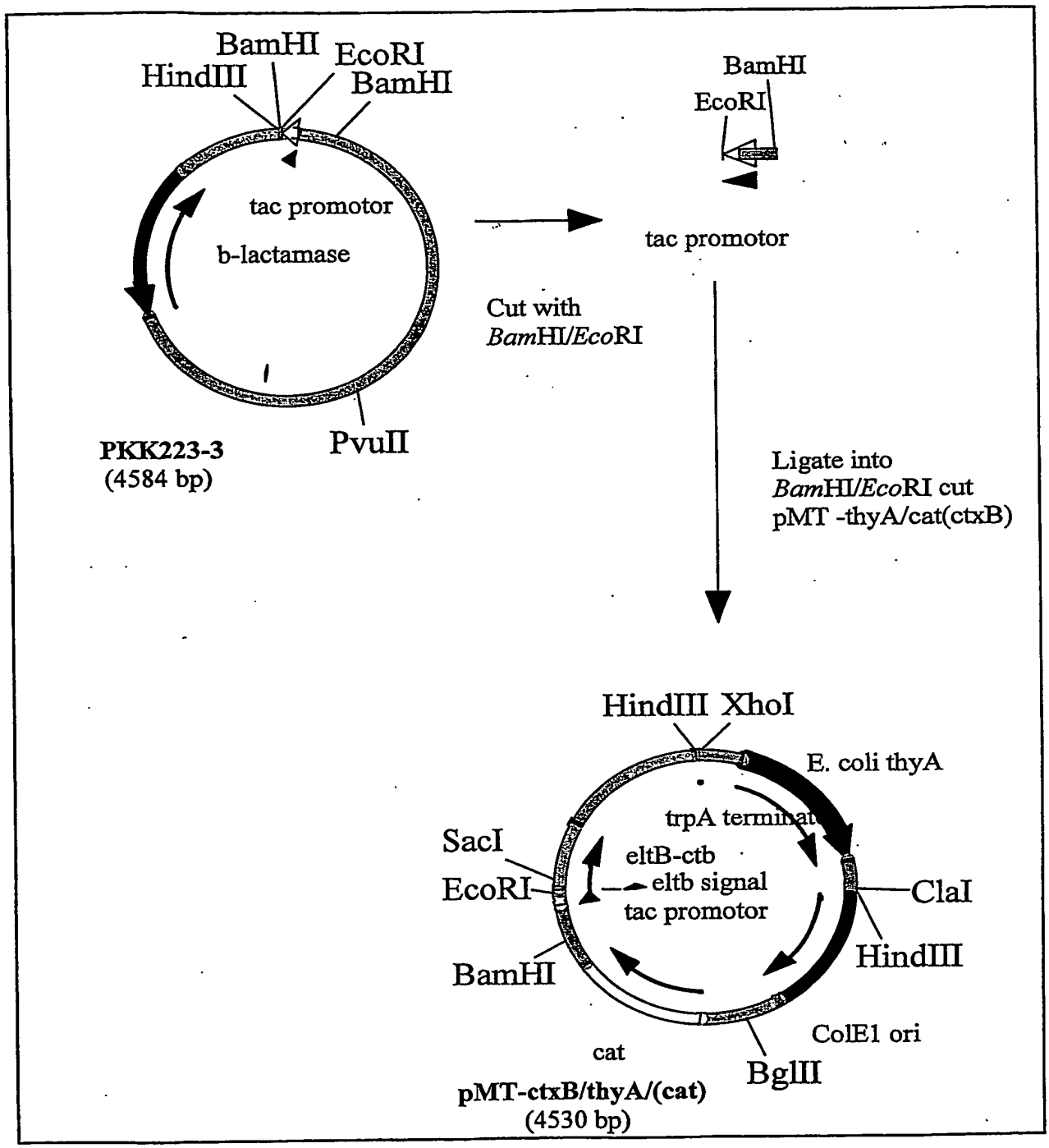

BamHI
EcoRI
HindIII
BamHI
tac promotor
b-lactamase
PvuII
PKK223-3
(4584 bp)
Cut with
*Bam*HI/*Eco*RI
BamHI
EcoRI
tac promotor
Ligate into
*Bam*HI/*Eco*RI cut
pMT -thyA/cat(ctxB)
HindIII XhoI
E. coli thyA
trpA terminat
SacI
EcoRI
eltB-ctb
eltb signal
tac promotor
ClaI
BamHI
HindIII
ColE1 ori
cat
BglII
pMT-ctxB/thyA/(cat)
(4530 bp)

Figure 12

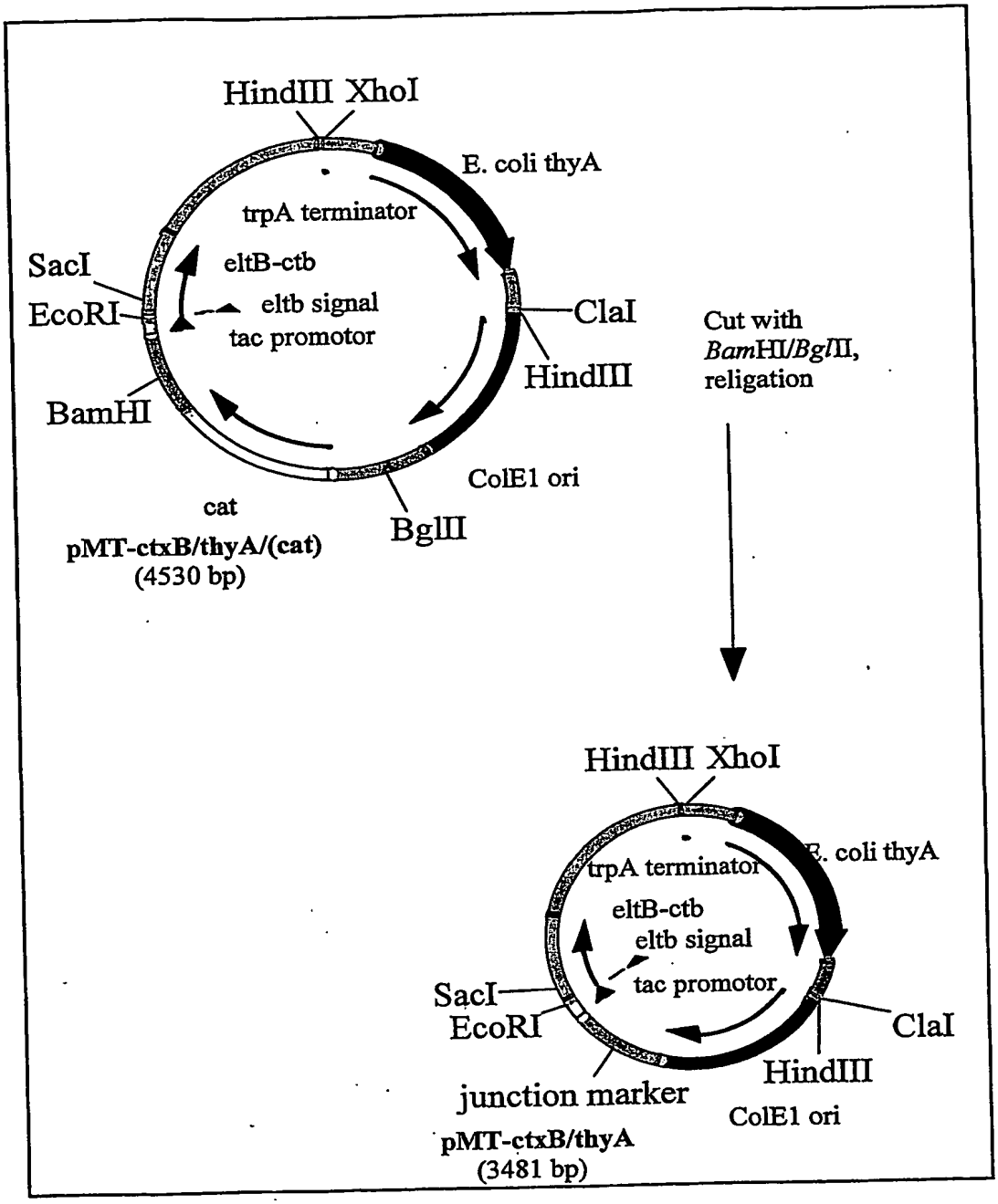

HindIII XhoI
E. coli thyA
trpA terminator
SacI
EcoRI
eltB-ctb
eltb signal
tac promotor
ClaI
HindIII
BamHI
ColE1 ori
cat
BglII
pMT-ctxB/thyA/(cat)
(4530 bp)
Cut with
BamHI/BglII,
religation
HindIII XhoI
E. coli thyA
trpA terminator
eltB-ctb
eltb signal
tac promotor
SacI
EcoRI
ClaI
HindIII
junction marker
ColE1 ori
pMT-ctxB/thyA
(3481 bp)

Figure 13

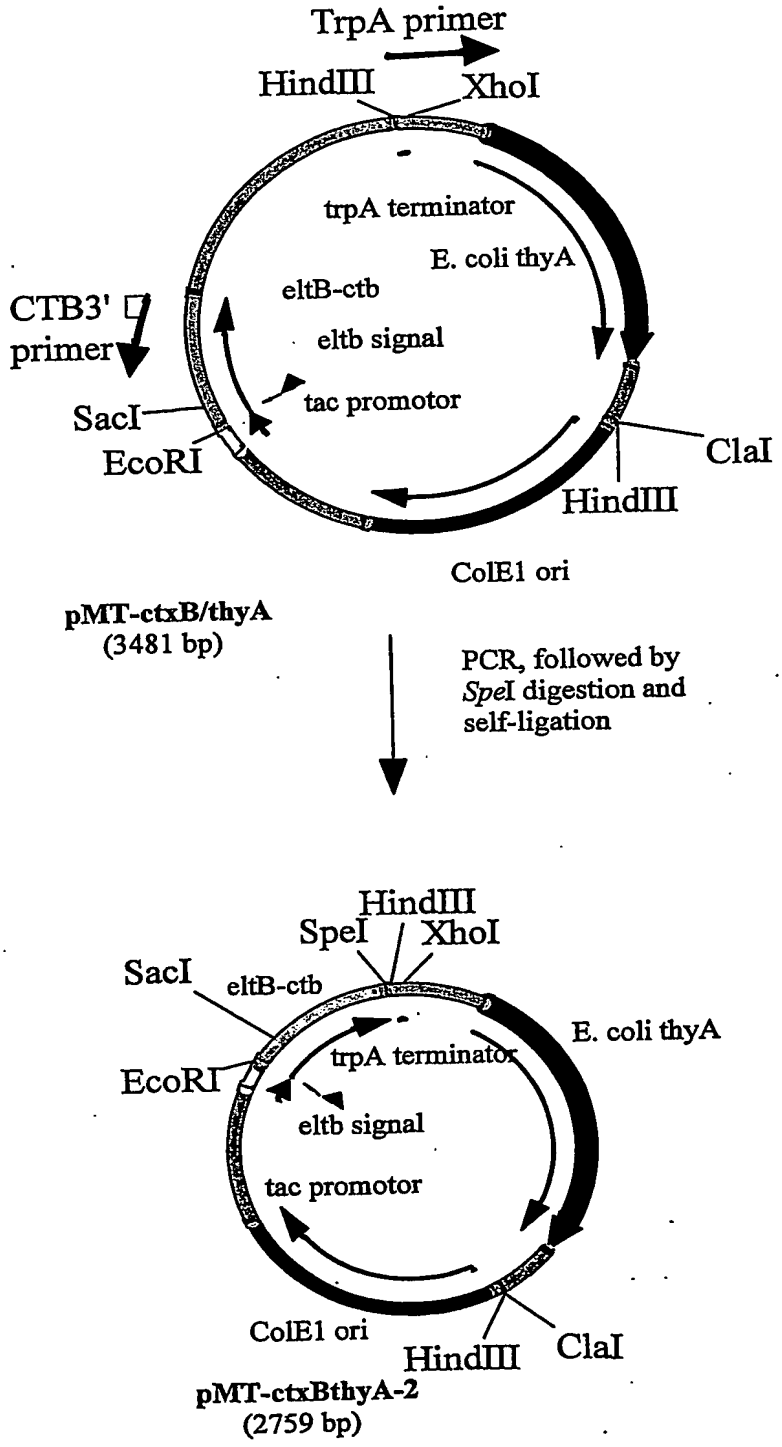
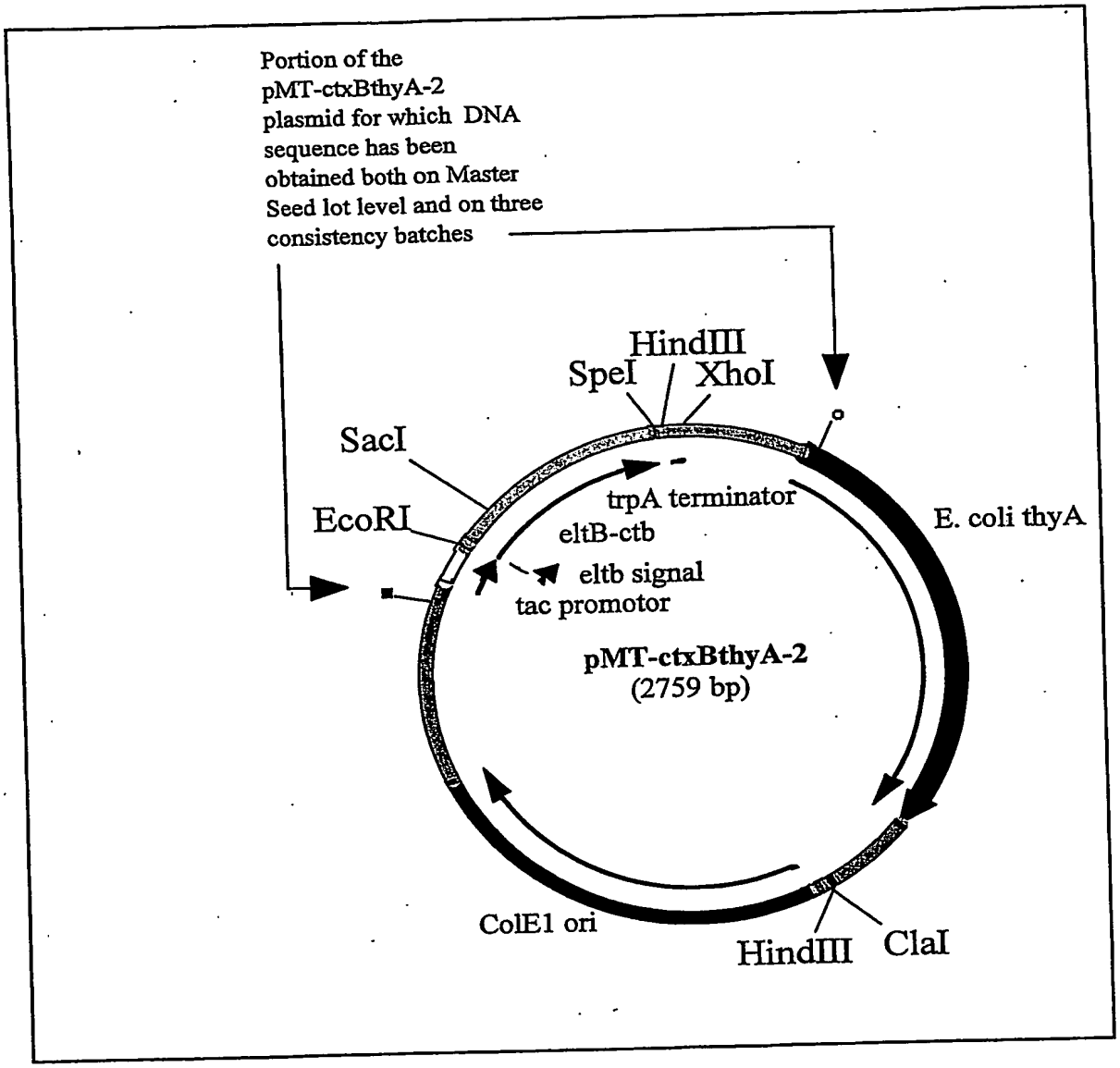

TrpA primer
HindIII
XhoI
trpA terminator
E. coli thyA
eltB-ctb
eltb signal
CTB3' primer
tac promotor
SacI
EcoRI
ClaI
HindIII
ColE1 ori
pMT-ctxB/thyA
(3481 bp)
PCR, followed by
SpeI digestion and
self-ligation
HindIII
SpeI
XhoI
SacI
eltB-ctb
E. coli thyA
EcoRI
trpA terminator
eltb signal
tac promotor
ColE1 ori
HindIII
ClaI
pMT-ctxBthyA-2
(2759 bp)

Figure 14


Portion of the
pMT-ctxBthyA-2
plasmid for which DNA
sequence has been
obtained both on Master
Seed lot level and on three
consistency batches
HindIII
SpeI
XhoI
SacI
EcoRI
trpA terminator
eltB-ctb
eltb signal
tac promotor
E. coli thyA
pMT-ctxBthyA-2
(2759 bp)
ColE1 ori
HindIII
ClaI

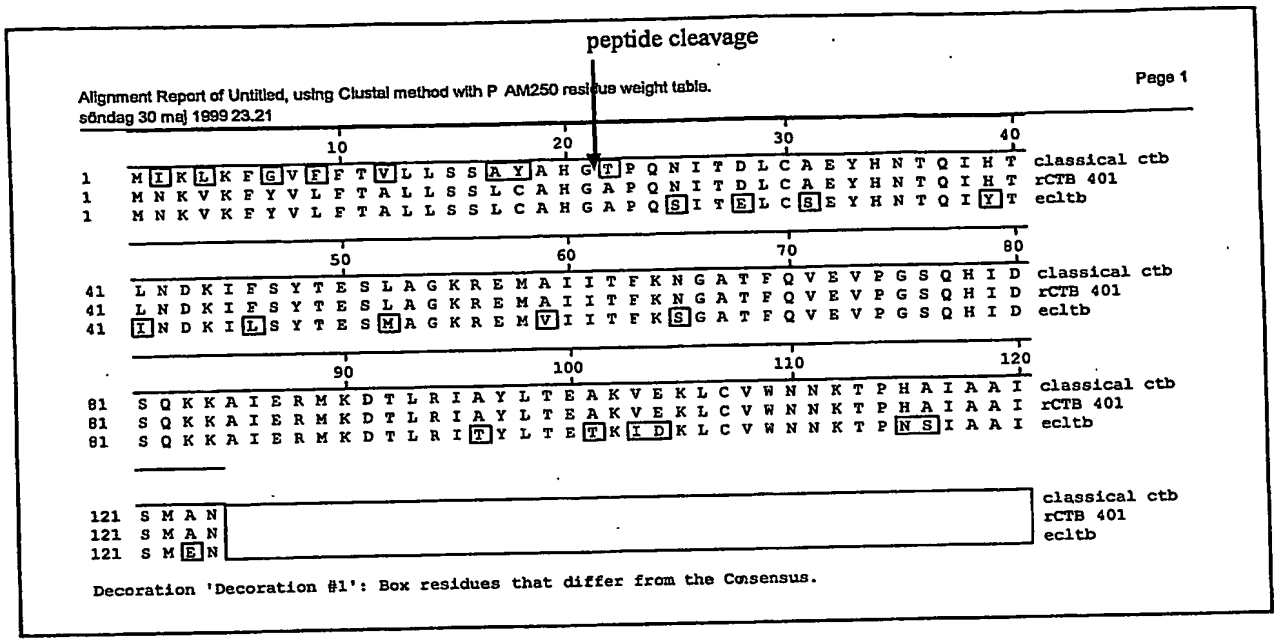

peptide cleavage
Alignment Report of Untitled, using Clustal method with P AM250 residue weight table.
Page 1
söndag 30 maj 1999 23.21
10 20 30 40
1 M I K L K F G V F F T V L L S S A Y A H G T P Q N I T D L C A E Y H N T Q I H T classical ctb
1 M N K V K F Y V L F T A L L S S L C A H G A P Q N I T D L C A E Y H N T Q I H T rCTB 401
1 M N K V K F Y V L F T A L L S S L C A H G A P Q S I T E L C S E Y H N T Q I Y T ecltb
50 60 70 80
41 L N D K I F S Y T E S L A G K R E M A I I T F K N G A T F Q V E V P G S Q H I D classical ctb
41 L N D K I F S Y T E S L A G K R E M A I I T F K N G A T F Q V E V P G S Q H I D rCTB 401
41 I N D K I L S Y T E S M A G K R E M V I I T F K S G A T F Q V E V P G S Q H I D ecltb
90 100 110 120
81 S Q K K A I E R M K D T L R I A Y L T E A K V E K L C V W N N K T P H A I A A I classical ctb
81 S Q K K A I E R M K D T L R I A Y L T E A K V E K L C V W N N K T P H A I A A I rCTB 401
81 S Q K K A I E R M K D T L R I T Y L T E T K I D K L C V W N N K T P N S I A A I ecltb
121 S M A N classical ctb
121 S M A N rCTB 401
121 S M E N ecltb
Decoration 'Decoration #1': Box residues that differ from the Consensus.

Figure 15

Figure 16

```
          |  10      |  20      |  30      |  40      |  50      |  60      |  70      |  80      |  90      | 100
    1 CTCGAGGTTT GTTCCTGATT GGTTACGGCG CGTTTCGCAT CATTGTTGAG  TTTTTCCGCC AGCCCGACGC GCAGTTTACC GGTGCCTGGG TGCAGTACAT 100

  101 CAGCATGGGG CAAATTCTTT CCATCCCGAT GATTGTCGCG GGTGTGATCA  TGATGGTCTG GGCATATCGT CGCAGCCCAC AGCAACACGT TTCCTGAGGA 200
        start thyA
  201 ACCATGAAAC AGTATTTAGA ACTGATGCAA AAAGTGCTCG ACGAAGGCAC  ACAGAAAAAC GACCGTACCG GAACCGGAAC GCTTTCCATT TTTGGTCATC 300

  301 AGATGCGTTT TAACCTGCAA GATGGATTCC GGCTGGTGAC AACTAAACGT  TGCCACCTGC GTTCCATCAT CCATGAACTG CTGTGGTTTC TGCAGGGCGA 400

  401 CACTAACATT GCTTATCTAC ACGAAAACAA TGTCACCATC TGGGACGAAT  GGGCCGATGA AAACGGCGAC CTCGGGCCAG TGTATGGTAA ACAGTGGCGC 500

  501 GCCTGGCCAA CGCCAGATGG TCGTCATATT GACCAGATCA CTACGGTACT  GAACCAGCTG AAAAACGACC CGGATTCGCG CCGCATTATT GTTTCAGCGT 600

  601 GGAACGTAGG CGAACTGGAT AAAATGGCGC TGGCACCGTG CCATGCATTC  TTCCAGTTCT ATGTGGCAGA CGGCAAACTC TCTTGCCAGC TTTATCAGCG 700

  701 CTCCTGTGAC GTCTTCCTCG GCCTGCCGTT CAACATTGCC AGCTACGCGT  TATTGGTGCA TATGATGGCG CAGCAGTGCG ATCTGGAAGT GGGTGATTTT 800

  801 GTCTGGACCG GTGGCGACAC GCATCTGTAC AGCAACCATA TGGATCAAAC  TCATCTGCAA TTAAGCCGCG AACCGCGTCC GCTGCCGAAG TTGATTATCA 900
                                                                                                  end thyA
  901 AACGTAAACC CGAATCCATC TTCGACTACC GTTTCGAAGA CTTTGAGATT  GAAGGCTACG ATCCGCATCC GGGCATTAAA GCGCCGGTGG CTATCTAATT 1000

 1001 ACGAAACATC CTGCCAGAGC CGACGCCAGT GTGCGTCGGT TTTTTTACCC  TCCGTTAAAT TCTTCGAGAC GCCTTCCCGA AATTTTGCAA CGTCCTGCAA 1100

 1101 CGGCGTAAAT AGTCCGGAAG ATGCGCCGAA GAAATAGAAA CGTCGAATCA  AGCTTATCGA TACCGTCGAC CTTGAAGAGC AAGGATCTAG GTGAAGATCC 1200

 1201 TTTTTGATAA TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC  TGAGCGTCAG ACCCCGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTTT 1300

 1301 TTTTCTGCGC GTAATCTGCT GCTTGCAAAC AAAAAAACCA CCGCTACCAG  CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT TCCGAAGGTA 1400

 1401 ACTGGCTTCA GCAGAGCGCA GATACCAAAT ACTGTCCTTC TAGTGTAGCC  GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG 1500

 1501 CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT  CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC 1600

 1601 GGGCTGAACG GGGGGTTCGT GCACACAGCC CAGCTTGGAG CGAACGACCT  ACACCGAACT GAGATACCTA CAGCGTGAGC TATGAGAAAG CGCCACGCTT 1700
```

Figure 16 continued

```
1701 CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC  AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG CATCTTTATA 1800

1801 GTCCTGTCGG GTTTCGCCAC CTCTGACTTG ATCGTCGATT TTTGTGATGC  TCGTCAGGGG GGCGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT 1900

1901 ACGGTTCCTG GCCTTTTGCT GGCCTTTTGC TCACATGTTC TTTCCTGCGT  TATCCCCTGA TTCTGTGGAT AACCGTATTA CCGCCTTTGA GTGAGCTGAT 2000

2001 ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG TGAGCGAGGA  AGCGATGGAA GAGCAGATCC GGGCTTATCG ACTGCACGGT GCACCAATGC 2100

2101 TTCTGGCGTC AGGCAGCCAT CGGAAGCTGT GGTATGGCTG TGCAGGTCGT  AAATCACTGC ATAATTCGTG TCGCTCAAGG CGCACTCCCG TTCTGGATAA 2200

2201 TGTTTTTTGC GCCGACATCA TAACGGTTCT GGCAAATATT CTGAAATGAG  CTGTTGACAA TTAATCATCG GCTCGTATAA TGTGTGGAAT TGTGAGCGGA 2300
                                    start eltB
2301 TAACAATTTC ACACAGGAAA CAGAATTCGG GATGAATTAT GAATAAAGTA  AAATTTTATG TTTTATTTAC GGCGTTACTA TCCTCTCTAT GTGCACACGG 2400
start ctxB
2401 AGCTCCTCAA AATATTACTG ATTTGTGTGC AGAATACCAC AACACACAAA  TACATACGCT AAATGATAAG ATATTTTCGT ATACAGAATC TCTAGCTGGA 2500

2501 AAAAGAGAGA TGGCTATCAT TACTTTTAAG AATGGTGCAA CTTTTCAAGT  AGAAGTACCA GGTAGTCAAC ATATAGATTC ACAAAAAAAA GCGATTGAAA 2600

2601 GGATGAAGGA TACCCTGAGG ATTGCATATC TTACTGAAGC TAAAGTCGAA  AAGTTATGTG TATGGAATAA TAAAACGCCT CATGCGATTG CCGCAATTAG 2700
       end ctxB                          trpA terminator                                                          2759
2701 TATGGCAAAT TAAACTAGTC AATTGAAGCT TAGCCCGCCT AATGAGCGGG  CTTTTTTTT
         |  10     |  20     |  30     |  40     |  50      |  60     |  70     |  80     |  90     |  100
```

204-295: *E. coli thyA* coding region
1192-1876: Col E1 origin of replication
2339-2710: *eltB-ctxB* coding region
2402-2710: *ctxB* coding region
2732-2759: *trpA* terminator

Figure 17

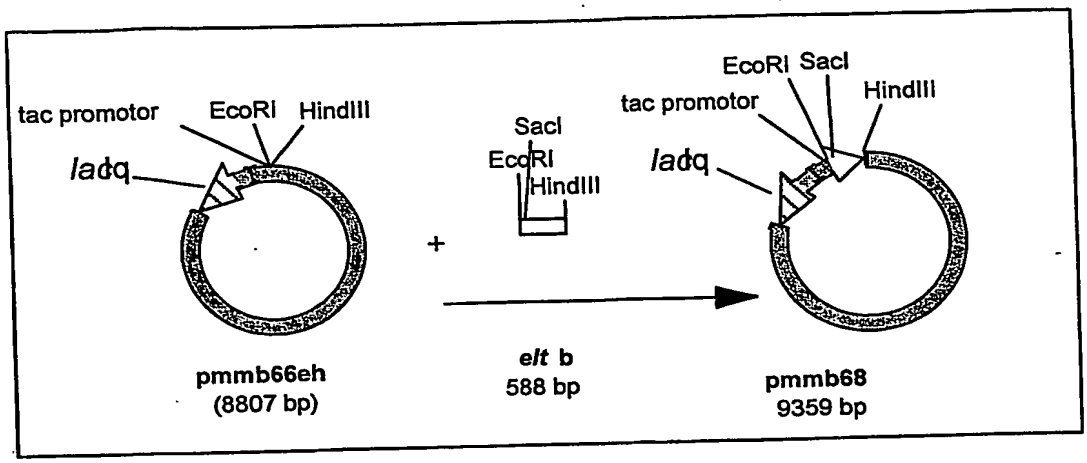

tac promotor
EcoRI
HindIII
lacIq
pmmb66eh
(8807 bp)
+
SacI
EcoRI
HindIII
elt b
588 bp
EcoRI
SacI
HindIII
tac promotor
lacIq
pmmb68
9359 bp

Figure 18

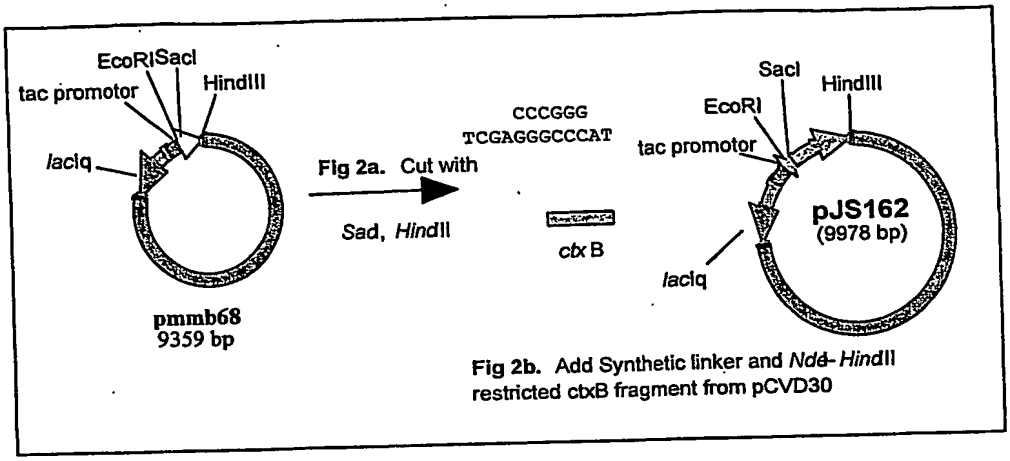


Fig 2b. Add Synthetic linker and *NdeI-HindII* restricted ctxB fragment from pCVD30

Figure 19

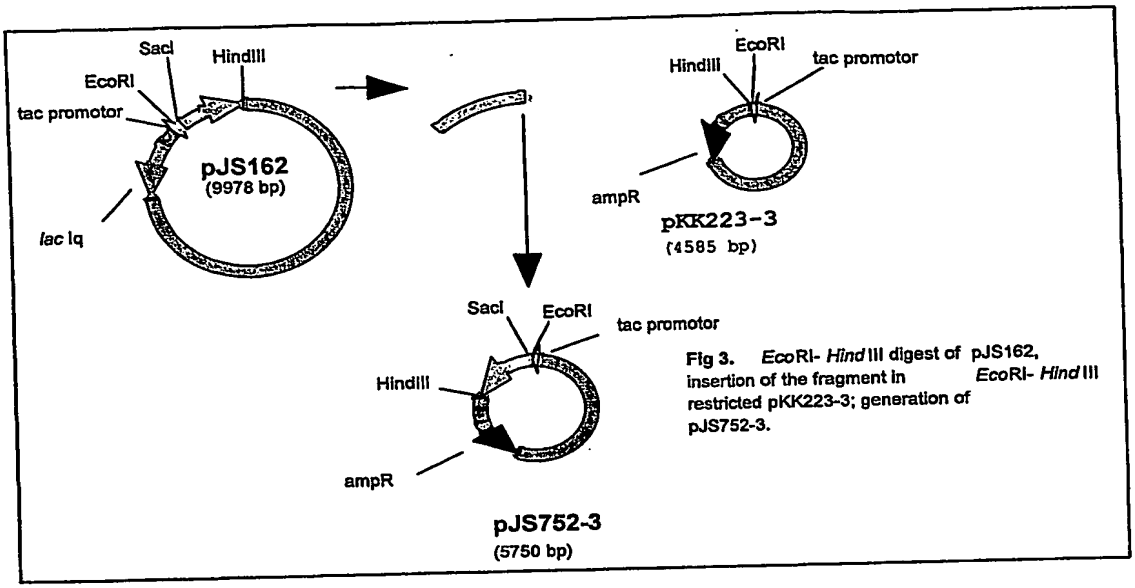


Fig 3. *Eco*RI- *Hind* III digest of pJS162, insertion of the fragment in *Eco*RI- *Hind* III restricted pKK223-3; generation of pJS752-3.

Figure 20

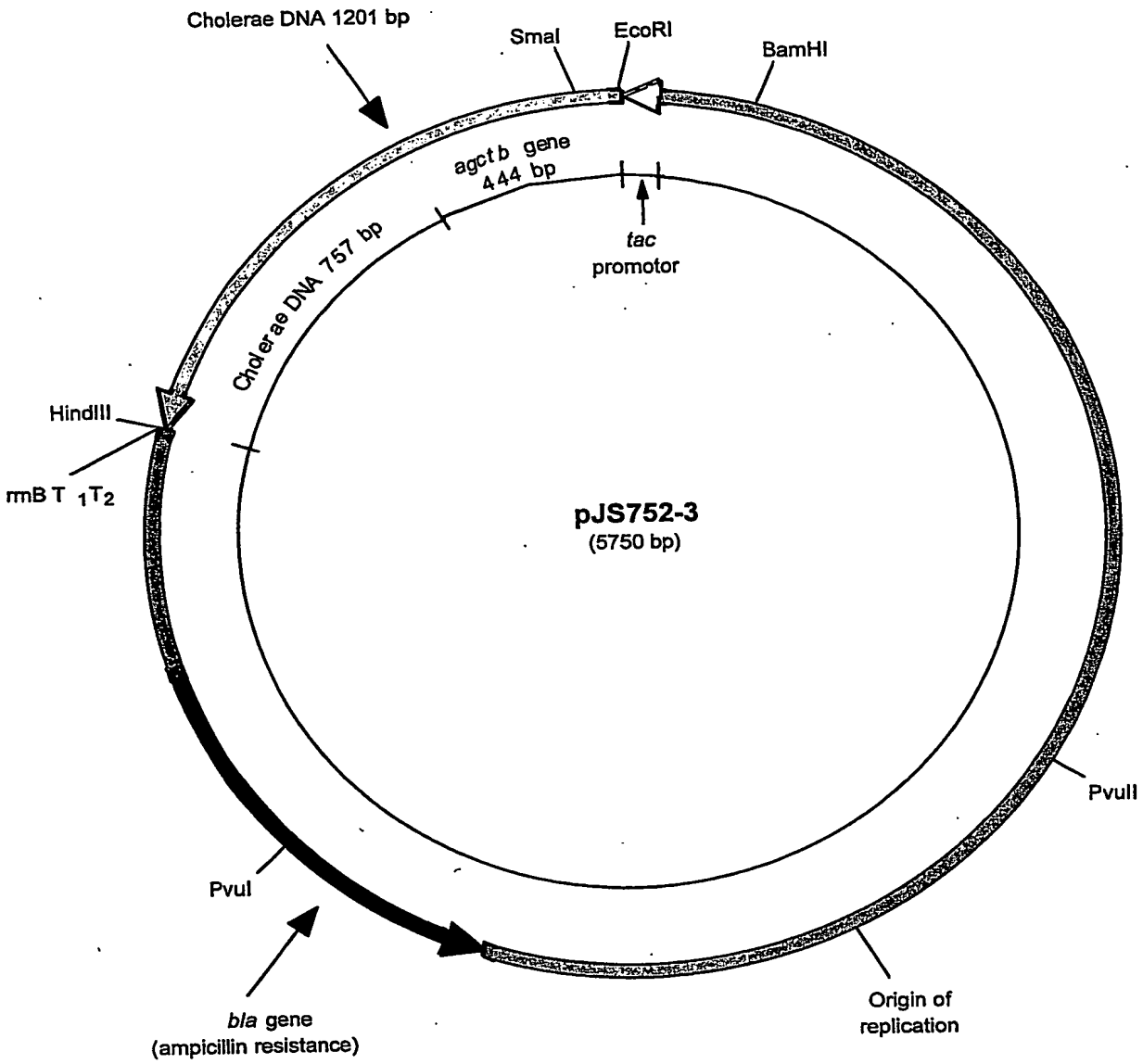

Cholerae DNA 1201 bp
SmaI
EcoRI
BamHI
agctb gene
444 bp
tac
promotor
Cholerae DNA 757 bp
HindIII
rrnB T 1T2
pJS752-3
(5750 bp)
PvuII
PvuI
bla gene
(ampicillin resistance)
Origin of
replication

Figure 21

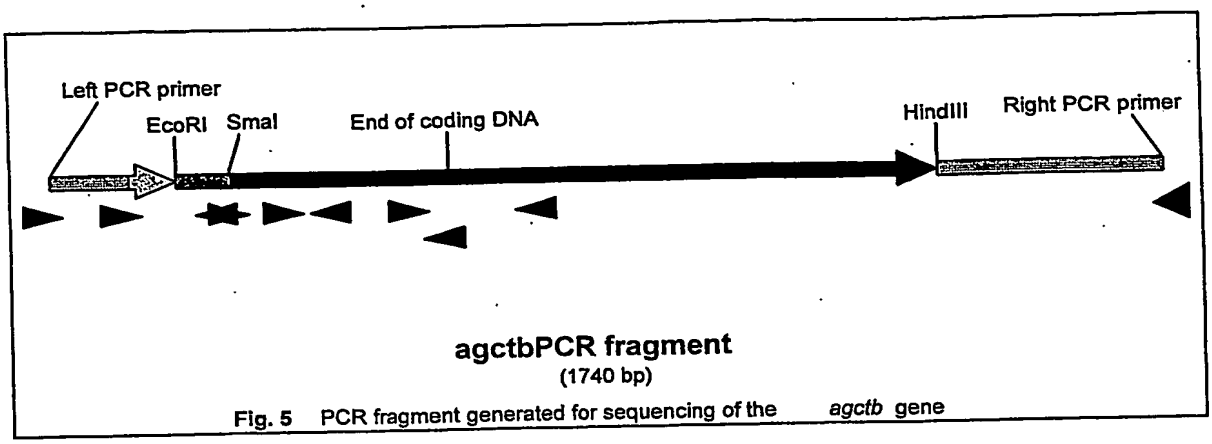


Fig. 5 PCR fragment generated for sequencing of the *agctb* gene

Figure 22

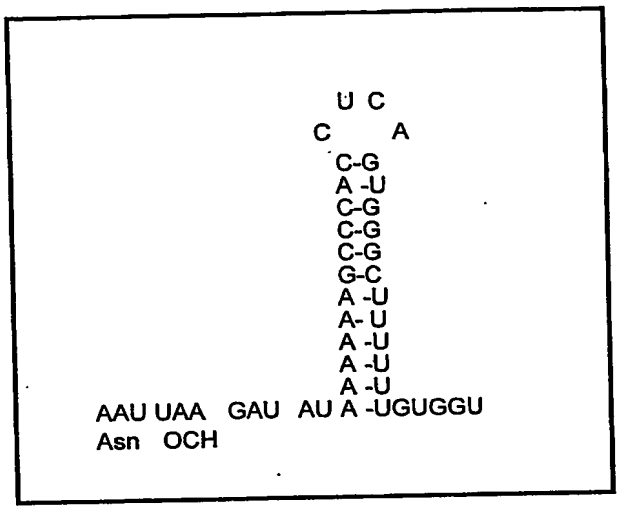

U C
C A
C-G
A -U
C-G
C-G
C-G
G-C
A -U
A- U
A -U
A -U
A -U
AAU UAA GAU AU A -UGUGGU
Asn OCH

Figure 23

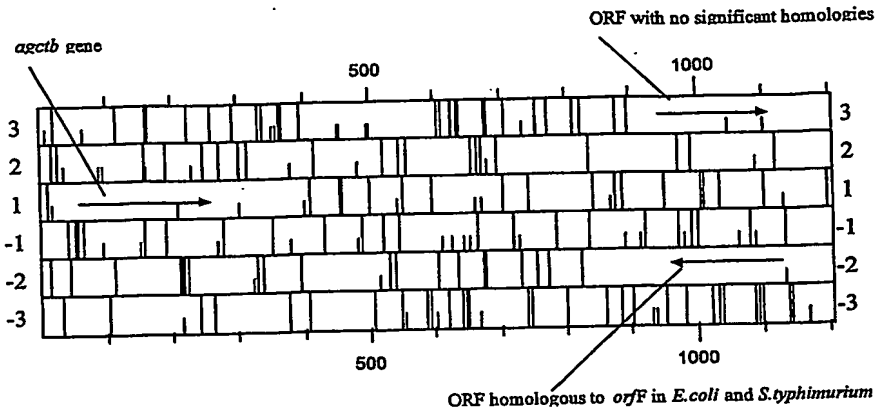

ORF with no significant homologies
agctb gene
500
1000
3
2
1
-1
-2
-3
3
2
1
-1
-2
-3
500
1000
ORF homologous to orfF in E.coli and S.typhimurium

Figure 24

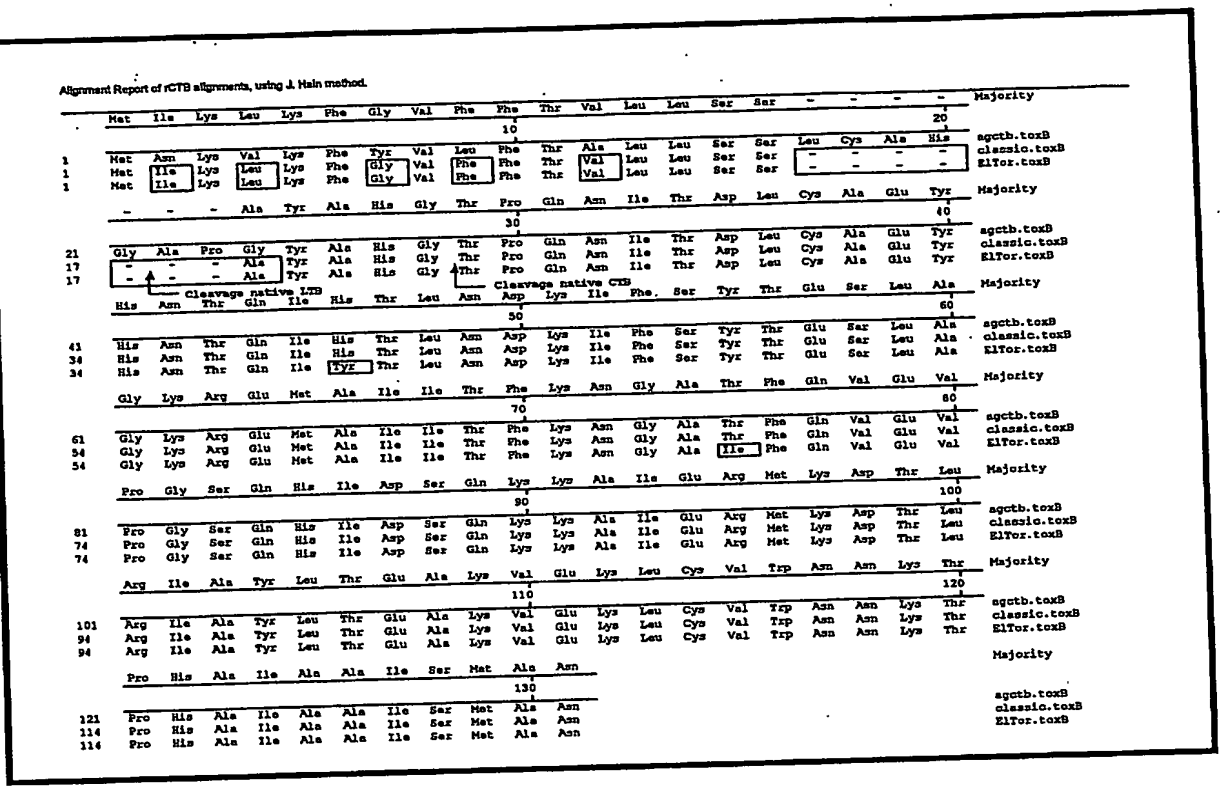

Alignment Report of rCTB alignments, using J. Hein method.
Majority
agctb.toxB
classic.toxB
ElTor.toxB
Cleavage native LTB
Cleavage native CTB

Figure 25

```
   | 10   | 20   | 30   | 40   | 50   | 60   | 70   | 80
1 AATTCGGGAT GAATTATGAA TAAAGTAAAA TTTTATGTTT TATTTACGGC GTTACTATCC TCTCTATGTG CACACGGAGC

81 TCCCGGGTAT GCACATGGAA CACCTCAAAA TATTACTGAT TTGTGTGCAG AATACCACAA CACACAAATA CATACGCTAA

161 ATGATAAGAT ATTTTCGTAT ACAGAATCTC TAGCTGGAAA AAGAGAGATG GCTATCATTA CTTTTAAGAA TGGTGCAACT

241 TTTCAAGTAG AAGTACCAGG TAGTCAACAT ATAGATTCAC AAAAAAAAGC GATTGAAAGG ATGAAGGATA CCCTGAGGAT

321 TGCATATCTT ACTGAAGCTA AAGTCGAAAA GTTATGTGTA TGGAATAATA AAACGCCTCA TGCGATTGCC GCAATTAGTA

401 TGGCAAATTA AGATATAAAA AAGCCCACCT CAGTGGGCTT TTTTGTGGTT CGATGATGAG AAGCAACCGT TTTGCCCAAA

481 CATGTATTAC TGCAAGTATG ATGTTTTTAT TCCACATCCT TAGTGCGTAT TACGTGCGTT ATGTTAAATT AAGGCATAAA

561 AAGAGGTCGC AAACCCCAAT CTGCTCCCTA TTCTTAATCC AGCATTAAAC ATAACTGTAT TTACCATAAA ATACCTAATA

641 ATCATATTTA TCATTTGACA ATGTTAAGCA TGATTATTAA TGACAGTGTT TGTGTAATCC CTTGATTTGA AAACCTGTTA

721 CCAACACATT TCAAATGTGT TTAGGTTTAC TTCGCTAAAA AAGAACTCCA CACTAAGGTG GAGTTATTGG CAGTTTGGGG

801 TTACACGAGG AGTAAAAATAC GACTTAACCA CCTGCTAATT TGACTTTGTG TCCTTTCGAT TCAAGCAATG TTTTGAGTTG

881 ATCTCTAACA TCTCCTTGAA TTTCAATATC GCCATCTTTT ACAGCTCCAC CACATCCACA TTTTTTCTTG AGTTCAGCGG

961 CAAGCAGCTT TAGAGCGGCA TCGTCAAGGT CTAAACCTTT TACAATAGTG ACACCTTTTC CTTTACGGCC TTTAGTTTCA

1041 CGGAAAATGC GGACAATACC ATCGCCTTTT GGACGCTCTA TTTTTTCATG CTCAGGCTTA ATGCGCCCGA CTTCGGTTGA

1121 ATACACCAAT GTCATACTTT TACTATCTTT TTTGTTGCGC TTGTTGTTGC ATCGCTCGTT GTCTTGCAAC GAGGTAAGCT

1201 TGGCTGTTTT GGCGGATGAG AGAAGATTTT CAGCCTGATA CAGATTAAAT CAGAACGCAG AAGCGGTCTG ATAAAACAGA
```

Figure 25 continued

1281 ATTTGCCTGG CGGCAGTAGC GCGGTGGTCC CACCTGACCC CATGCCGAAC TCAGAAGTGA AACGCCGTAG CGCCGATGGT

1361 AGTGTGGGGT CTCCCCATGC GAGAGTAGGG AACTGCCAGG CATCAAATAA AACGAAAGGC TCAGTCGAAA GACTGGGCCT

1441 TTCGTTTTAT CTGTTGTTTG TCGGTGAACG CTCTCCTGAG TAGGACAAAT CCGCCGGGAG CGGATTTGAA CGTTGCGAAG

1521 CAACGGCCCG GAGGGTGGCG GGCAGGACGC CCGCCATAAA CTGCCAGGCA TCAAATTAAG CAGAAGGCCA TCCTGACGGA

1601 TCCTGACGGA TGGCCTTTTT GCGTTTCTAC AAACTCTTTT GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT

1681 CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTTT

1761 TGCGGCATTT TGCCTTCCTG TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC

1841 GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG

1921 AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTGTTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA

2001 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT

2081 GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC

2161 GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA

2241 CGAGCGTGAC ACCACGATGC TGTAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA CTCTAGCTTC

2321 CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT

10 | 20 | 30 | 40 | 50 | 60 | 70 | 80

Figure 25 continued

```
      | 10   | 20   | 30   | 40   | 50   | 60   | 70   | 80
2401 TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC

2481 CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC

2561 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTTA

2641 AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA CTGAGCGTCA

2721 GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC

2801 ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC

2881 AGATACCAAA TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC

2961 GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT

3041 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC TACACCGAAC

3121 TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC

3201 AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA

3281 CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT

3361 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT

3441 ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA

3521 GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT ATTTCACACC GCATATGGTG CACTCTCAGT ACAATCTGCT

3601 CTGATGCCGC ATAGTTAAGC CAGTATACAC TCCGCTATCG CTACGTGACT GGGTCATGGC TGCGCCCCGA CACCCGCCAA

3681 CACCCGCTGA CGCGCCCTGA CGGGCTTGTC TGCTCCCGGC ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC
```

Figure 25 continued

3761 ATGTGTCAGA GGTTTTCACC GTCATCACCG AAACGCGCGA GGCAGCTGCG GTAAAGCTCA TCAGCGTGGT CGTGAAGCGA

3841 TTCACAGATG TCTGCCTGTT CATCCGCGTC CAGCTCGTTG AGTTTCTCCA GAAGCGTTAA TGTCTGGCTT CTGATAAAGC

3921 GGGCCATGTT AAGGGCGGTT TTTTCCTGTT TGGTCACTGA TGCCTCCGTG TAAGGGGGAT TTCTGTTCAT GGGGGTAATG

4001 ATACCGATGA AACGAGAGAG GATGCTCACG ATACGGGTTA CTGATGATGA ACATGCCCGG TTACTGGAAC GTTGTGAGGG

4081 TAAACAACTG GCGGTATGGA TGCGGCGGGA CCAGAGAAAA ATCACTCAGG GTCAATGCCA GCGCTTCGTT AATACAGATG

4161 TAGGTGTTCC ACAGGGTAGC CAGCAGCATC CTGCGATGCA GATCCGGAAC ATAATGGTGC AGGGCGCTGA CTTCCGCGTT

4241 TCCAGACTTT ACGAAACACG GAAACCGAAG ACCATTCATG TTGTTGCTCA GGTCGCAGAC GTTTTGCAGC AGCAGTCGCT

4321 TCACGTTCGC TCGCGTATCG GTGATTCATT CTGCTAACCA GTAAGGCAAC CCCGCCAGCC TAGCCGGGTC CTCAACGACA

4401 GGAGCACGAT CATGCGCACC CGTGGCCAGG ACCCAACGCT GCCCGAGATG CGCCGCGTGC GGCTGCTGGA GATGGCGGAC

4481 GCGATGGATA TGTTCTGCCA AGGGTTGGTT TGCGCATTCA CAGTTCTCCG CAAGAATTGA TTGGCTCCAA TTCTTGGAGT

4561 GGTGAATCCG TTAGCGAGGT GCCGCCGGCT TCCATTCAGG TCGAGGTGGC CCGGCTCCAT GCACCGCGAC GCAACGCGGG

4641 GAGGCAGACA AGGTATAGGG CGGCGCCTAC AATCCATGCC AACCCGTTCC ATGTGCTCGC CGAGGCGGCA TAAATCGCCG

4721 TGACGATCAG CGGTCCAGTG ATCGAAGTTA GGCTGGTAAG AGCCGCGAGC GATCCTTGAA GCTGTCCCTG ATGGTCGTCA

4801 TCTACCTGCC TGGACAGCAT GGCCTGCAAC GCGGGCATCC CGATGCCGCC GGAAGCGAGA AGAATCATAA TGGGGAAGGC

4881 CATCCAGCCT CGCGTCGCGA ACGCCAGCAA GACGTAGCCC AGCGCGTCGG CCGCCATGCC GGCGATAATG GCCTGCTTCT

4961 CGCCGAAACG TTTGGTGGCG GGACCAGTGA CGAAGGCTTG AGCGAGGGCG TGCAAGATTC CGAATACCGC AAGCGACAGG

| 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80

Figure 25 continued

| 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80

5041 CCGATCATCG TCGCGCTCCA GCGAAAGCGG TCCTCGCCGA AAATGACCCA GAGCGCTGCC GGCACCTGTC CTACGAGTTG

5121 CATGATAAAG AAGACAGTCA TAAGTGCGGC GACGATAGTC ATGCCCCGCG CCCACCGGAA GGAGCTGACT GGGTTGAAGG

5201 CTCTCAAGGG CATCGGTCGA CGCTCTCCCT TATGCGACTC CTGCATTAGG AAGCAGCCCA GTAGTAGGTT GAGGCCGTTG

5281 AGCACCGCCG CCGCAAGGAA TGGTGCATGC AAGGAGATGG CGCCCAACAG TCCCCCGGCC ACGGGGCCTG CCACCATACC

5361 CACGCCGAAA CAAGCGCTCA TGAGCCCGAA GTGGCGAGCC CGATCTTCCC CATCGGTGAT GTCGGCGATA TAGGCGCCAG

5441 CAACCGCACC TGTGGCGCCG GTGATGCCGG CCACGATGCG TCCGGCGTAG AGGATCCGGG CTTATCGACT GCACGGTGCA

5521 CCAATGCTTC TGGCGTCAGG CAGCCATCGG AAGCTGTGGT ATGGCTGTGC AGGTCGTAAA TCACTGCATA ATTCGTGTCG

5601 CTCAAGGCGC ACTCCCGTTC TGGATAATGT TTTTTGCGCC GACATCATAA CGGTTCTGGC AAATATTCTG AAATGAGCTG

5681 TTGACAATTA ATCATCGGCT CGTATAATGT GTGGAATTGT GAGCGGATAA CAATTTCACA CAGGAAACAG

| 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80

Underline = sequenced in master seed lots and consistency batches.

Bold and italic = Open reading frame of Eltb leader sequence and ctxB gene = agctb gene

Figure 26

Alignment Report of rCTB + master alignments, using J. Hein method with Weighted residue weight table.

tisdag 12 augusti 1997 15:38

```
Majority        G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G
                          10        20        30        40

      1550      G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  121
      1551      G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  124
      1552      G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G   98
      1553      G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  124
      1554      G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  122
      1555      G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  118
Master No 1     G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  230
Master No 2     G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T C T G A A A T G A G  272

Majority        C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T
                          50        60        70        80

      1550      C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  161
      1551      C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  164
      1552      C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  138
      1553      C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  164
      1554      C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  162
      1555      C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  158
Master No 1     C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  270
Master No 2     C T G T T G A C A A T T A A T C A T C G G C T C G T A T A A T G T G T G G A A T  312

Majority        T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G
                          90        100       110       120

      1550      T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  201
      1551      T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  204
      1552      T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  178
      1553      T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  204
      1554      T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  202
      1555      T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  198
Master No 1     T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  310
Master No 2     T G T G A G C G G A T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  352

Majority        G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C
                          130       140       150       160

      1550      G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  241
      1551      G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  244
      1552      G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  218
      1553      G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  244
      1554      G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  242
      1555      G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  238
Master No 1     G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  350
Master No 2     G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G T T T T A T T T A C  392

Majority        G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G
                          170       180       190       200

      1550      G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  281
      1551      G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  284
      1552      G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  258
      1553      G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  284
      1554      G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  282
      1555      G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  278
Master No 1     G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  390
Master No 2     G G C G T T A C T A T C C T C T C T A T G T G C A C A C G G A G C T C C C G G G  432

Majority        T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G
                          210       220       230       240

      1550      T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  321
      1551      T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  324
      1552      T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  298
      1553      T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  324
      1554      T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  322
      1555      T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  318
Master No 1     T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  430
Master No 2     T A T G C A C A T G G A A C A C C T C A A A A T A T T A C T G A T T T G T G T G  472
```

Figure 26. continued ...

Alignment Report of rCTB + master alignments, using J. Hein method with Weighted residue weight table.
tisdag 12 augusti 1997 15:48

```
Majority        C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A
                                  250       260       270       280
      1550      C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  361
      1551      C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  364
      1552      C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  338
      1553      C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  364
      1554      C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  362
      1555      C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  358
Master No 1     C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  470
Master No 2     C A G A A T A C C A C A A C A C A C A A A T A C A T A C G C T A A A T G A T A A  512

Majority        G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G
                                  290       300       310       320
      1550      G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  401
      1551      G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  404
      1552      G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  378
      1553      G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  404
      1554      G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  402
      1555      G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  398
Master No 1     G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  510
Master No 2     G A T A T T T T C G T A T A C A G A A T C T C T A G C T G G A A A A A G A G A G  552

Majority        A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G
                                  330       340       350       360
      1550      A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  441
      1551      A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  444
      1552      A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  418
      1553      A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  444
      1554      A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  442
      1555      A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  438
Master No 1     A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  550
Master No 2     A T G G C T A T C A T T A C T T T T A A G A A T G G T G C A A C T T T T C A A G  592

Majority        T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A
                                  370       380       390       400
      1550      T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  481
      1551      T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  484
      1552      T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  458
      1553      T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  484
      1554      T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  482
      1555      T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  478
Master No 1     T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  590
Master No 2     T A G A A G T A C C A G G T A G T C A A C A T A T A G A T T C A C A A A A A A A  632

Majority        A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T
                                  410       420       430       440
      1550      A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  521
      1551      A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  524
      1552      A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  498
      1553      A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  524
      1554      A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  522
      1555      A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  518
Master No 1     A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  630
Master No 2     A G C G A T T G A A A G G A T G A A G G A T A C C C T G A G G A T T G C A T A T  672

Majority        C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A
                                  450       460       470       480
      1550      C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  561
      1551      C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  564
      1552      C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  538
      1553      C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  564
      1554      C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  562
      1555      C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  558
Master No 1     C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  670
Master No 2     C T T A C T G A A G C T A A A G T C G A A A A G T T A T G T G T A T G G A A T A  712
```

Figure 26 continued ...

Alignment Report of rCTB + master alignments, using J. Hein method with Weighted residue weight table.

fisdag 12 augusti 1997 15:49

```
Majority  A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A
                                  490       500       510       520

1550      A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  601
1551      A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  604
1552      A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  578
1553      A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  604
1554      A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  602
1555      A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  598
Master No 1 A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  710
Master No 2 A T A A A A C G C C T C A T G C G A T T G C C G C A A T T A G T A T G G C A A A A  752

Majority  T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G
                            530       540       550       560

1550      T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  641
1551      T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  644
1552      T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  618
1553      T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  644
1554      T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  642
1555      T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  638
Master No 1 T T A A G A T A T A A A A A A A G C C C A C C T C A G T G G G C T T T T T T G T G  750
Master No 2 T T A A G A T A T A A A A A A A G C C C A C C . T C A G T G G G C T T T T T T G T G  792

Majority  G T T C G A T G A T G A G A A G C A
                        570

1550      G T T C G A T G A T G A G A A G C A  659
1551      G T T C G A T G A T G A G A A G C A  662
1552      G T T C G A T G A T G A G A A G C A  636
1553      G T T C G A T G A T G A G A A G C A  662
1554      G T T C G A T G A T G A G A A G C A  660
1555      G T T C G A T G A T G A G A A G C A  656
Master No 1 G T T C G A T G A T G A G A A G C A  768
Master No 2 G T T C G A T G A T G A G A A G C A  810
```

Figure 27

Alignment Report of Untitled, using J. Hein method with Weighted residue weight table.

onsdag 26 maj 1999 16.28

```
                  360                 370                 380
358  G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T  pMT-ctxBthyA-2
355  G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T  Master tube 1
371  G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T  Master tube 2
375  G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T  M4806
376  G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T  M4807
378  G C C G A C A T C A T A A C G G T T C T G G C A A A T A T T  M4808

                  390                 400                 410
388  C T G A A A T G A G C T G T T G A C A A T T A A T C A T C G  pMT-ctxBthyA-2
385  C T G A A A T G A G C T G T T G A C A A T T A A T C A T C G  Master tube 1
401  C T G A A A T G A G C T G T T G A C A A T T A A T C A T C G  Master tube 2
405  C T G A A A T G A G C T G T T G A C A A T T A A T C A T C G  M4806
406  C T G A A A T G A G C T G T T G A C A A T T A A T C A T C G  M4807
408  C T G A A A T G A G C T G T T G A C A A T T A A T C A T C G  M4808

                  420                 430                 440
418  G C T C G T A T A A T G T G T G G A A T T G T G A G C G G A  pMT-ctxBthyA-2
415  G C T C G T A T A A T G T G T G G A A T T G T G A G C G G A  Master tube 1
431  G C T C G T A T A A T G T G T G G A A T T G T G A G C G G A  Master tube 2
435  G C T C G T A T A A T G T G T G G A A T T G T G A G C G G A  M4806
436  G C T C G T A T A A T G T G T G G A A T T G T G A G C G G A  M4807
438  G C T C G T A T A A T G T G T G G A A T T G T G A G C G G A  M4808

                  450                 460                 470
448  T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  pMT-ctxBthyA-2
445  T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  Master tube 1
461  T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  Master tube 2
465  T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  M4806
466  T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  M4807
468  T A A C A A T T T C A C A C A G G A A A C A G A A T T C G G  M4808

                  480                 490                 500
478  G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G  pMT-ctxBthyA-2
475  G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G  Master tube 1
491  G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G  Master tube 2
495  G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G  M4806
496  G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G  M4807
498  G A T G A A T T A T G A A T A A A G T A A A A T T T T A T G  M4808

                  510                 520                 530
508  T T T T A T T T A C G G C G T T A C T A T C C T C T C T A T  pMT-ctxBthyA-2
505  T T T T A T T T A C G G C G T T A C T A T C C T C T C T A T  Master tube 1
521  T T T T A T T T A C G G C G T T A C T A T C C T C T C T A T  Master tube 2
525  T T T T A T T T A C G G C G T T A C T A T C C T C T C T A T  M4806
526  T T T T A T T T A C G G C G T T A C T A T C C T C T C T A T  M4807
528  T T T T A T T T A C G G C G T T A C T A T C C T C T C T A T  M4808

                  540                 550                 560
538  G T G C A C A C G G A G C T C C T C A A A A T A T T A C T G  pMT-ctxBthyA-2
535  G T G C A C A C G G A G C T C C T C A A A A T A T T A C T G  Master tube 1
551  G T G C A C A C G G A G C T C C T C A A A A T A T T A C T G  Master tube 2
555  G T G C A C A C G G A G C T C C T C A A A A T A T T A C T G  M4806
556  G T G C A C A C G G A G C T C C T C A A A A T A T T A C T G  M4807
558  G T G C A C A C G G A G C T C C T C A A A A T A T T A C T G  M4808

                  570                 580                 590
568  A T T T G T G T G C A G A A T A C C A C A A C A C A C A A A  pMT-ctxBthyA-2
565  A T T T G T G T G C A G A A T A C C A C A A C A C A C A A A  Master tube 1
581  A T T T G T G T G C A G A A T A C C A C A A C A C A C A A A  Master tube 2
585  A T T T G T G T G C A G A A T A C C A C A A C A C A C A A A  M4806
586  A T T T G T G T G C A G A A T A C C A C A A C A C A C A A A  M4807
588  A T T T G T G T G C A G A A T A C C A C A A C A C A C A A A  M4808
```

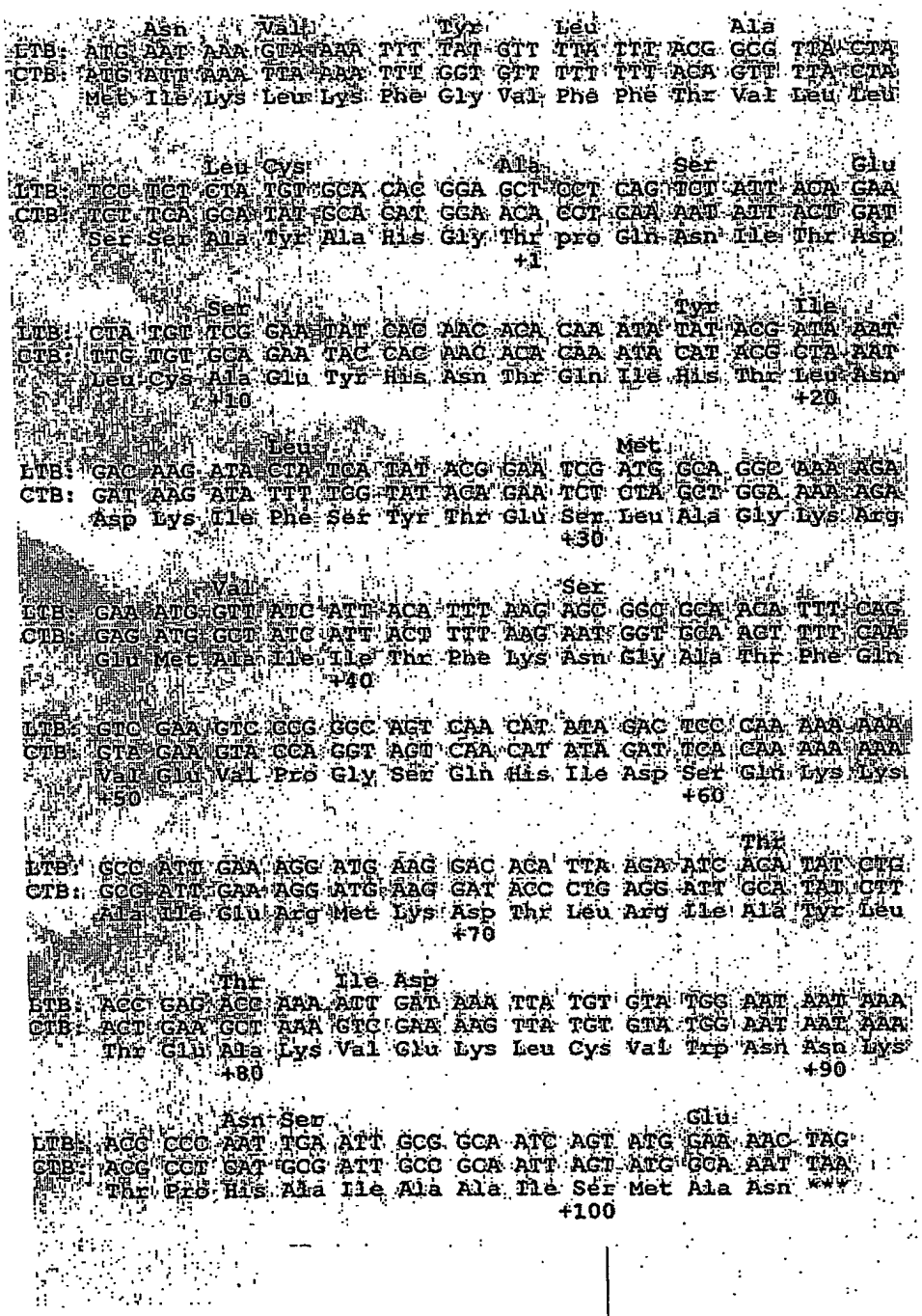

Asn Val Tyr Leu Ala
LTB: ATG AAT AAA GTA AAA TTT TAT GTT TTA TTT ACG GCG TTA CTA
CTB: ATG ATT AAA TTA AAA TTT GGT GTT TTT TTT ACA GTT TTA CTA
Met Ile Lys Leu Lys Phe Gly Val Phe Phe Thr Val Leu Leu
Leu Cys Ala Ser Glu
LTB: TCC TCT CTA TGT GCA CAC GGA GCT CCT CAG TCT ATT ACA GAA
CTB: TCT TCA GCA TAT GCA CAT GGA ACA CCT CAA AGT ATT ACT GAT
Ser Ser Ala Tyr Ala His Gly Thr pro Gln Asn Ile Thr Asp
+1
Ser Tyr Ile
LTB: CTA TGT TCG GAA TAT CAC AAC ACA CAA ATA TAT ACG ATA AAT
CTB: TTG TGT GCA GAA TAC CAC AAC ACA CAA ATA CAT ACG CTA AAT
Leu Cys Ala Glu Tyr His Asn Thr Gln Ile His Thr Leu Asn
+10 +20
Leu Met
LTB: GAC AAG ATA CTA TCA TAT ACG GAA TCG ATG GCA GGC AAA AGA
CTB: GAT AAG ATA TTT TCG TAT AGA GAA TCT CTA GCT GGA AAA AGA
Asp Lys Ile Phe Ser Tyr Thr Glu Ser Leu Ala Gly Lys Arg
+30
Val Ser
LTB: GAA ATG GTT ATC ATT ACA TTT AAG AGC GGC GCA ACA TTT CAG
CTB: GAG ATG GCT ATC ATT ACT TTT AAG AAT GGT GCA AGT TTT CAA
Glu Met Ala Ile Ile Thr Phe Lys Asn Gly Ala Thr Phe Gln
+40
LTB: GTC GAA GTC CCG GGC AGT CAA CAT ATA GAC TCC CAA AAA AAA
CTB: GTA GAA GTA CCA GGT AGT CAA CAT ATA GAT TCA CAA AAA AAA
Val Glu Val Pro Gly Ser Gln His Ile Asp Ser Gln Lys Lys
+50 +60
Thr
LTB: GCC ATT GAA AGG ATG AAG GAC ACA TTA AGA ATC ACA TAT CTG
CTB: GCG ATT GAA AGG ATG AAG GAT ACC CTG AGG ATT GCA TAT CTT
Ala Ile Glu Arg Met Lys Asp Thr Leu Arg Ile Ala Tyr Leu
+70
Thr Ile Asp
LTB: ACC GAG ACG AAA ATT GAT AAA TTA TGT GTA TGG AAT AAT AAA
CTB: ACT GAA GCT AAA GTC GAA AAG TTA TGT GTA TGG AAT AAT AAA
Thr Glu Ala Lys Val Glu Lys Leu Cys Val Trp Asn Asn Lys
+80 +90
Asn Ser Glu
LTB: ACG CCC AAT TCA ATT GCG GCA ATC AGT ATG GAA AAC TAG
CTB: ACG CCT GAT GCG ATT GCC GCA ATT AGT ATG GCA AAT TAA
Thr Pro His Ala Ile Ala Ala Ile Ser Met Ala Asn ***
+100

Figure 28